What Causal Forces Shape Internet Connectivity at the AS-level?

Hyunseok Chang
Department of EECS
University of Michigan
Ann Arbor, MI 48109-2122
hschang@eecs.umich.edu

Sugih Jamin
Department of EECS
University of Michigan
Ann Arbor, MI 48109-2122
jamin@eecs.umich.edu

Walter Willinger
AT&T Labs-Research
180 Park Ave.
Florham Park, NJ 07932-0971
walter@research.att.com

Abstract— **Two ASs are connected in the Internet AS graph only if they have a business "peering relationship." By focusing on the AS subgraph AS_{PC} whose links represent** *provider-customer* **relationships, we present an empirical study that identifies three crucial causal forces at work in the design of AS connectivity: (i)** *AS-geography***, i.e., locality and number of PoPs within individual ASs; (ii)** *AS-specific business models***, abstract toy models that describe how individual ASs choose their "best" provider; and (iii)** *AS evolution***, a historic account of the "lives" of individual ASs in a dynamic ISP market. Based on these findings that directly relate to how provider-customer relationships may be determined in the actual Internet, we develop a new optimization-driven model for Internet growth at the AS_{PC} level. Its defining feature is an explicit construction of a novel class of intuitive, multi-objective, local optimizations by which the different ASs determine in a fully distributed and decentralized fashion their "best" upstream provider. We show that our model is broadly robust, perforce yields graphs that match inferred AS connectivity with respect to many different metrics, and is ideal for exploring the impact of new peering incentives or policies on AS-level connectivity.**

I. INTRODUCTION

Internet connectivity at the level of Autonomous Systems (ASs) reflects existing business relationships among ASs. Two ASs are connected in an AS graph by a link only if they have a "peering relationship" between them, e.g., *provider-customer* or *peer-to-peer* relationship. In principle, snapshots of the Internet's AS graph can be inferred from BGP-derived measurements, but in practice, the resulting graph structures require careful interpretation. For example, since the measurements consist of a collection of snapshots of BGP routing tables taken at a few vantage points on the Internet over time, private peering links and backup connections between ASs generally cannot be identified, because they remain largely invisible to collection sites such as the Oregon route server [28]. Thus, the resulting inferred AS graphs tend to be less densely connected than the actual Internet at the AS level [12], and the inferred peering relationships are not always accurate. These difficulties not withstanding, a striking characteristic of various inferred AS graphs (with more or less incomplete connectivity information) has been the observed high variability of the AS vertex degrees, parsimoniously captured by vertex degree distributions of the power-law type [15], [12]. This power-law finding has led to renewed interest in network modeling and has motivated the development of new topology generators. However, much of the efforts to date has been either abstract (e.g., generate a graph with a given vertex degree distribution [1], [21]) or based on some exogenously imposed mechanisms (e.g., a presumed preferential-type connectivity rule [2], [24]). While the resulting models and generators are generally successful in reproducing and matching the power-law type node degree distributions of measured AS graphs, their relevance to networking is seriously hampered by their generic nature—they are mostly designed to model all types of networks that show power-law type node degree distributions [7]. As a result, these models completely ignore any AS-specific factors and criteria inherent in establishing the very business relationships expressed by AS graphs. The models' theoretical appeal is also limited, mainly because of the models' exclusive focus on a single metric (i.e., node degree distribution) and a general inability to also match in a parsimonious manner inferred AS graphs with respect to alternative metrics (e.g., hierarchy-related or graph evolution-specific measures).

The main objective of this paper is to explore a radically different approach to modeling and generating Internet topologies at the AS level.[1] Instead of relying on

This project is funded in part by NSF grant number ANI-0082287 and by ONR grant number N000140110617. Sugih Jamin is further supported by the NSF CAREER Award ANI-9734145, the Presidential Early Career Award for Scientists and Engineers (PECASE) 1998, and the Alfred P. Sloan Foundation Research Fellowship 2001. Additional funding is provided by AT&T Research, and by equipment grants from Sun Microsystems Inc. and Compaq Corp.

[1] A similar approach has recently been advocated and outlined in [3] for modeling Internet connectivity at the router-level, but no results have been reported to date.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2003**	2. REPORT TYPE	3. DATES COVERED **00-00-2003 to 00-00-2003**
4. TITLE AND SUBTITLE **What Causal Forces Shape Internet Connectivity at the AS-level?**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Michigan,Computer Science and Engineering Division,Department of Electrical Engineering and Computer Science,Ann Arbor,MI,48109-2122**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S)
		11. SPONSOR/MONITOR’S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**23**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

abstract or artificially imposed mechanisms that yield AS graphs which manage to attain highly variable node degree distributions but fail to match other graph characteristics, we focus here on identifying key causal forces at work in the fully distributed and decentralized design of AS interconnectivity. More specifically, we are mainly concerned with the subgraph AS_{PC} of the overall AS graph whose links represent only provider-customer peering relationships, i.e., non provider-customer relationships are not part of this subgraph. We show that by explicitly incorporating provider-customer business relationships (or concrete abstractions thereof) into an appropriate network growth model, the resulting AS subgraphs automatically match inferred AS_{PC}-specific connectivity with respect to a wide range of metrics. Furthermore, these subgraphs share a number of characteristics with the overall AS graphs (e.g., highly variable node degree distributions). This finding demonstrates the importance of the AS_{PC} subgraphs in gaining a better understanding of the Internet's overall AS connectivity and it argues for a detailed study of the properties and dynamic nature of these AS_{PC} subgraphs.

Our overall approach is motivated by the recently proposed HOT (*Highly Optimized Tolerance*) concept introduced by Carlson and Doyle [11], [13]. HOT provides a general framework in which highly variable event sizes, in systems highly optimized by engineering design, are the result of tradeoffs between yield, cost of resources, and the systems' tolerance to risk. In turn, the HOT framework emphasizes the importance of design, structure, and optimization in the study of highly engineered and complex systems, such as the Internet. In applying this general HOT concept to the specific problem addressed in this paper, our work draws heavily on the first explicit attempt by Fabrikant *et al.* [14] to cast network design as a HOT problem.[2] In Fabrikant *et al.*'s generic toy model of network growth, each newly arriving node decides on the particular node of the existing graph it connects to by solving a generic multi-objective optimization problem. This model serves as starting point of our investigation into the causal forces shaping Internet connectivity at the AS_{PC} level. That is, following Fabrikant *et al.*'s model, we attempt to determine a particular class of multi-objective optimization problems that reflect the various factors and criteria by which provider and customer ASs determine their peering relationships with one another.

Our proposed approach to Internet modeling at the AS_{PC} level has three major novel components. First, in contrast to the *generic* Internet growth model of Fabrikant *et al.* that deals with *generic* nodes and links, we provide the first explicit attempt to cast the design of AS level connectivity as a HOT problem. In particular, we develop a new HOT model for AS_{PC} graphs that fully exploits the networking semantics of the relevant objects—the graph's nodes represent businesses (i.e., provider or customer ISPs) and its links express explicit business relationships among these nodes. Second, we formulate a novel class of multi-objective optimization problems to capture the decision process ASs use to determine their peering or business relationships with other ASs. To arrive at this formulation, we proceed in three steps and show successively that (i) AS geography, (ii) AS business model, and (iii) AS evolution define three key criteria that must be accounted for. We propose concrete abstractions of each of these criteria and incorporate them one-by-one into our multi-objective optimization problems. We illustrate how these optimization problems shape the formation of local provider-customer relationships and cause the "emerging" global characteristics exhibited by the resulting AS_{PC} graphs. In addition, we show that a number of features associated with the overall AS graphs (e.g., high variability in node degree distributions) are already present in the corresponding AS_{PC} subgraphs and can therefore be explained and understood in terms of the causal forces at work in the design of the AS_{PC} portion of the overall AS graph. Finally, we illustrate along the way that the proposed HOT model for Internet growth at the AS_{PC} level has a number of attractive robustness properties with respect to the details of how the three identified key criteria are abstracted and expressed as objectives that have to be optimized simultaneously. These insensitivity results enhance the overall credibility of the proposed HOT model. They also make the resulting model especially appealing for exploring a range of what-if scenarios. For example, if modifications to existing criteria for establishing provider-customer relationships, or the introduction of new ones, may more accurately reflect incentives in future provider-customer relationships, our model can be used to explore the impact such modifications or embellishments may have on the overall AS connectivity.

Due to the generic nature of previous efforts to modeling Internet topologies, hitherto none of the three key criteria we propose has played any significant role in modeling the AS graphs. For the non-generic, networking-centric approach we propose, the importance of these three criteria should come as no surprise. Consider for example "AS geography," by which we mean the number and geographic locations of an AS's *Points-of-Presence* (PoPs). Clearly, knowing the geography of intra-AS PoP structures is important to, say, distinguish between nearby and far-away providers—with obvious implications in determining con-

[2] Fabrikant *et al.* also suggested *Heuristically Optimized Tradeoff* as a fitting alternative acronym for HOT.

nection costs. Similarly, depending on an AS's business model (which may include technological aspects such as network availability, network reliability, expected performance, or support for value-added services, as well as economic considerations such as pricing plans, projected network build-out, customer support, etc.), customer ASs can be expected to behave in a more or less rational manner when establishing upstream connectivity; for example, they might choose among competing providers by implicitly or explicitly satisfying some local Pareto optimality criterion with respect to some underlying utility measure. As for AS evolution, the dynamics of the ISP market (overall growth amidst ever-present mergers, acquisitions, and bankruptcies) cause or force ASs to periodically re-examine their peering relationships and, if necessary, establish new, get rid of old, or modify existing relationships—with obvious implications on the overall connectivity at the AS level.

We hasten to point out that while the generation of realistic AS topologies is also a natural by-product of our approach, it is *not* the focus of this paper. In particular, to extend our approach and develop a HOT model for Internet growth at the overall AS level will require incorporating peer-to-peer relationships and accounting for ASs that change from peer to provider/customer or vice versa. In this sense, the causal forces identified in this paper and accounted for in the proposed HOT model are necessary but not sufficient for shaping the existing Internet topology at the overall AS level. For example, while our model succeeds in explaining the power-law type node degree distributions of inferred AS graphs in terms of AS geography-related characteristics (i.e., AS size measured in terms of number of PoPs per AS), it also suggests the presence of additional causal forces that relate directly to the peer-to-peer portion of AS graphs. These forces will require careful consideration when extending our approach and developing a more complete model for Internet connectivity at the overall AS level. However, we leave the pursuit of such a modeling effort for future research.

The rest of the paper is structured as follows. In Section II, we describe the two inferred AS graphs that form the basis of our empirical study and reduce them to simpler tree topologies that are used in the subsequent sections. We discuss in Section III the HOT model for generic Internet growth proposed in [14] and demonstrate its shortcomings when applied to the AS graph. The construction of our HOT model for Internet growth at the AS_{PC} level is described in Sections IV–VI, where we identify, in succession, AS geography, AS business model, and AS evolution as key forces at work in the design of Internet AS connectivity. We conclude the paper with a discussion of some practical implication of our proposed approach and the ensuing models and comment on the work's impact on network modeling as a scientific discipline.

TABLE I

AS RELATIONSHIP INFERENCE RESULT

AS graph	Number of links			
	All	Provider-customer	Peer-to-peer	Others
Oregon	23,449	21,473 (91.6%)	1,621 (6.9%)	355
Oregon+	32,759	27,815 (84.9%)	3,919 (12.0%)	1,025

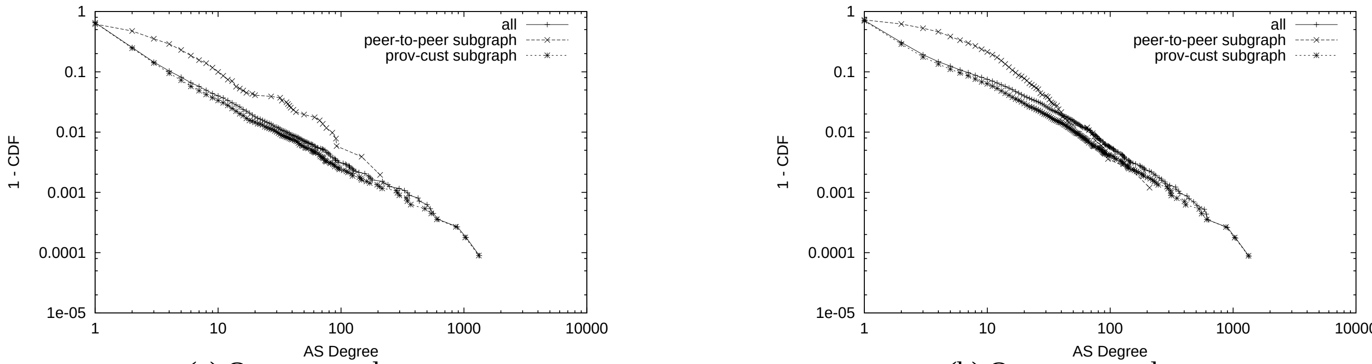
II. INFERENCE FOR AS GRAPHS

We address in this section some of the challenges related to inferring AS graphs and describe our approach for tackling them in ways that are necessarily imperfect, but still ensure the validity of our results.

A. Inferring AS Connectivity

In principle, to determine whether AS X and AS Y have a peering relationship with one another, all that is needed is to collect X's BGP routing table and check if AS Y appears in any of the table's AS-path entries. In practice, however, only a very limited number of ASs make their BGP routing tables publicly available. As a result, all available inferred AS graphs are necessarily incomplete. While this incompleteness should always be kept in mind when reporting observed characteristics of measured AS graphs, it does not necessarily invalidate all of the findings. A particularly attractive method to demonstrate the general validity of a given AS-related observation despite the incompleteness property is to illustrate that the result in question is robust with respect to alternative AS graph inference techniques or to reliance on additional relevant data (e.g., more BGP routing tables, Looking Glass information, etc.).

To explore the robustness of the findings reported in this paper, we rely on two inferred AS graphs that have been studied in detail in [12]. Most importantly, these AS graphs have significantly different connectivity densities. One graph, the OREGON AS graph, is constructed exclusively from information contained in the BGP routing tables collected by the Oregon route server [28]. This AS graph has been the most widely-used graph in past studies of the Internet's AS topology. The one used here is based on data collected in late May of 2001. The second graph, the OREGON+ AS graph, relies not only on the Oregon BGP data, but also makes use of a number of additional BGP routing tables, BGP summary information obtained by querying numerous ASs using the Looking Glass tool, and diligently extracted data from the Routing Registry



(a) Oregon topology (b) Oregon+ topology

Fig. 1. Node degree distributions of AS (sub)graphs.

TABLE II
AS GRAPH STATISTICS

AS graph		All	prov-cust subgraph All	prov-cust subgraph Largest comp.	peer-to-peer subgraph All	peer-to-peer subgraph Largest comp.
Oregon	# of nodes	11,183	11,177	11,177 (100%)	512	512 (100%)
	# of links	23,449	21,473	21,473 (100%)	1,621	1,621 (100%)
Oregon+	# of nodes	11,456	11,257	11,257 (100%)	1,365	837 (61.3%)
	# of links	32,759	27,815	27,815 (100%)	3,919	3,359 (85.7%)

database.[3] The data collection period is again late May 2001. While this graph has about the same number of ASs as the OREGON AS graph, Table II (first column) shows that it has about 40% more links than the OREGON AS graph. Realizing that both of these AS graphs exhibit qualitatively the same power-law type node degree distribution (e.g., see Fig. 1, lines labeled "all") suggests that this feature is robust with respect to the degree of incompleteness of the two graphs and can therefore be considered a genuine characteristic of Internet connectivity at the AS level. It also justifies our pursuit of a careful investigation into the underlying causes of this striking characteristic.

B. Provider-customer vs. Peer-to-peer Relationships

We are first interested in whether the power-law type node degree distributions observed in both of our AS graphs are associated with any specific peering type, in which case we may simplify our study by removing the non-causal link types from the graphs. A link between two ASs in an AS graph typically reflects either a *provider-customer* or a *peer-to-peer* relationship.[4] In the former, one AS plays the role of the customer, while the other is the provider of Internet connectivity. Internet providers are paid by their customers for providing this service. In the latter, the ASs see equal benefit in interconnecting with each other and no financial exchange takes place. Given the BGP routing tables (plus other information, if available) used to infer an AS graph, we can annotate the links in the graph with inferred peering relationships using heuristics proposed in [17] and [30].[5] The results of augmenting the OREGON and OREGON+ AS graphs with inferred peering relationships are summarized in Table I.

To determine whether the power-law type node degree distributions observed in our inferred AS graphs are associated with any specific peering type, we consider two separate subgraphs for each of the OREGON and OREGON+ AS graphs. The *provider-customer* (or prov-cust, for short) subgraph contains only provider-customer links (along with their incident ASs), while the *peer-to-peer* subgraph consists of peer-to-peer links only (and all their incident ASs).[6] Table II provides details on these subgraphs (or their largest connected components).

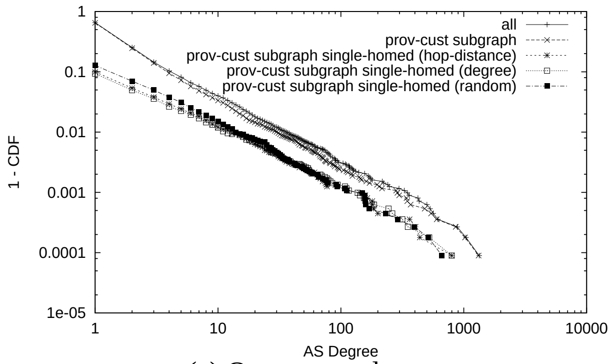
Fig. 1 compares the node degree distributions of the original OREGON AS graph (line labeled "all") and its two subgraphs (lines labeled "prov-cust subgraph" and "peer-to-peer subgraph") in plot (a), and the original OREGON+

[3]The use of selected information from the Routing Registry database is discussed and justified in [12] and is not essential for the purposes of this paper.

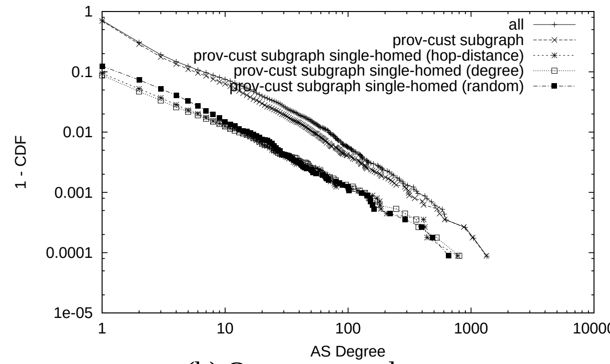
[4]Other relationships are possible, for example as backup links between non-provider ASs, but are relatively rare. They are grouped under "Others" in Table I.

[5]We need to use both heuristics due to limitations of each one of them. For example, the heuristic proposed in [17] cannot be used on our OREGON+ graph that contains links inferred from Looking Glass data but are not present in any of the collected BGP tables.

[6]In case one of the resulting subgraphs is not connected, we consider its largest connected component.



(a) Oregon topology



(b) Oregon+ topology

Fig. 2. Node degree distributions of P1 subgraphs.

AS graph and its two subgraphs in plot (b). As can be clearly seen, in both cases the node degree distributions of the provider-customer subgraphs are almost identical to those of the original AS graphs, but those corresponding to the peer-to-peer subgraphs behave qualitatively different. We can conclude that peer-to-peer and other non provider-customer links are non-causal factors as far as the power-law type node degree distribution characteristic of AS graphs is concerned. In turn, Fig. 1 suggests that it may suffice to focus mainly on the provider-customer portion of the AS graph when attempting to identify some of the causal forces at work in the design of Internet connectivity at the AS level.

C. Single- vs. Multi-homing

To further simplify our study of AS graph structures, we next ask if multi-homing has any major impact on the power-law type node degree distribution.[7] Multi-homing typically refers to a customer AS having more than one peering link with either the same provider AS or with different provider ASs. The purpose of multi-homing is mainly to improve an AS's reliability or performance (e.g., via load-balancing). To gauge the effect of multi-homing on AS-connectivity, we consider three heuristics for distinguishing between the *primary* connection and the *secondary* connection(s) of an AS.[8] For the first heuristic, given an AS X, we define its *best* provider (among its current providers) to be the AS with the lowest average hop distance to all the other ASs in the provider-customer subgraph. The second heuristic redefines the best provider of AS X to be the one with the largest AS node degree among all of X's current providers, while the third heuristic randomly picks one of AS X's providers as the best provider. Thus, by annotating the links of the provider-customer subgraph as *primary* (link to best provider) or *secondary*, we obtain three different instances of the subgraph, one per heuristic.

For each of these three annotated provider-customer subgraphs, we generate a new subgraph (called the P1 subgraph) by removing all secondary links. These "single-homed" P1 subgraphs of both the OREGON and OREGON+ provider-customer subgraphs have approximately 50% fewer links than their "multi-homed" counterparts, i.e., their corresponding provider-customer subgraphs. Fig. 2 depicts the node degree distributions of the inferred provider-customer subgraph (labeled "prov-cust subgraph") and its three P1 subgraphs (labeled "prov-cust subgraph single-homed (hop distance)," "prov-cust subgraph single-homed (degree)," and "prov-cust subgraph single-homed (random)") of the OREGON and OREGON+ graphs in plots (a) and (b), respectively. The figure shows that the P1 subgraphs exhibit qualitatively the same node degree distributions as the corresponding provider-customer subgraphs, except that the entire distributions are shifted down, with a slight change in slopes. Being largely independent of the original AS graph (i.e., OREGON vs. OREGON+) and of the details of defining a customer's "best" provider (i.e., the three instances of the P1 subgraph), Fig. 2 suggests that multi-homing is not responsible for the power-law type node degree distributions and not a causal force behind this property.

The P1 subgraphs, in which every AS has a single provider, form forest topologies, possibly with multiple root nodes. However, it turns out that for both the OREGON and OREGON+ graphs, the largest (connected) tree contained in their respective P1 subgraphs is rooted at AS701 (UUNET) and spans about 99% of all of P1's nodes. For the remainder of this paper, when exploring

[7] The increase in multi-homed networks has been a driving force behind the recent growth in the size of BGP routing tables [20].

[8] Multi-homing for the purpose of load-balancing may consider all the multi-homing links as equally important. However, in this paper, we assume that even in such a case, there still exists one primary upstream connection carrying a major portion of an AS's traffic.

the causal forces at work in shaping Internet connectivity at the AS_{PC} level, we will focus mainly on this simplified AS graph consisting of the largest connected tree topology rooted at AS701.

III. AS GROWTH—AN OPTIMIZATION-DRIVEN PROCESS

The orthodox physics views tend to associate the ubiquity of power-law distributions in natural and engineered complex systems unambiguously with critical phase transition [6]. However, in the specific case of the Internet, where power-law type distributions abound, this apparent connection turns out to be specious and can be directly refuted [32]. As a result, the Internet has become a prime target for testing the validity of alternative views and theories to explain the ubiquity of power-law type distributions in nature and engineering. An especially promising and radically different such alternative view has recently been proposed by Carlson and Doyle [11], [13] and is based on the concept of HOT (for *Highly Optimized Tolerance*). In this section, we summarize the state-of-the-art of using HOT-based approaches to model Internet growth and investigate their relevance for network design at the AS level.

A. The Generic HOT Model of Fabrikant et al.

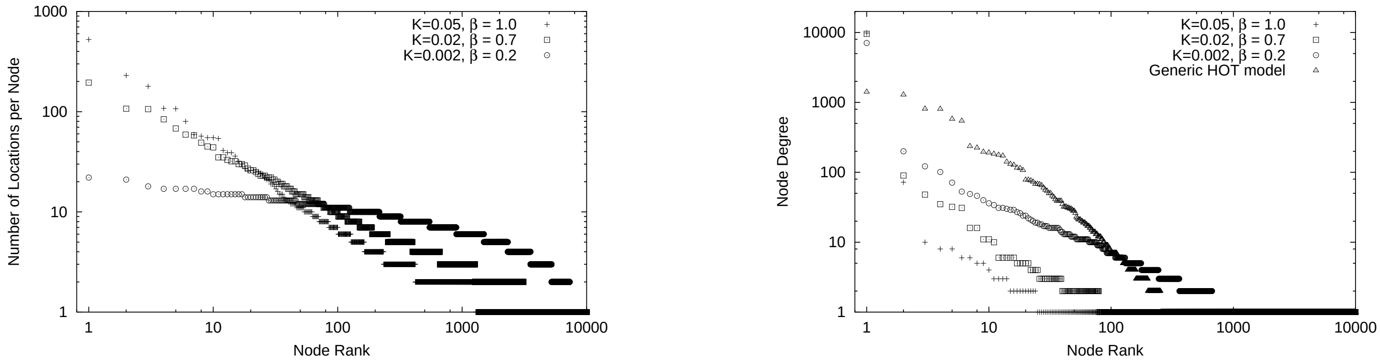
The HOT concept introduced in [11], [13] emphasizes the importance of design, structure, and optimization and provides a framework in which power-law type event size distributions in systems optimized by engineering design are the results of tradeoffs between yield, cost of resources, and tolerance to risk. *Highly Optimized* alludes to the fact that robustness (i.e., the maintenance of some desired system characteristics despite uncertainties in the behavior of its component parts or its environment) is achieved by highly structured, rare, non-generic configurations which—for highly engineered systems—are the result of deliberate design. *Tolerance* emphasizes that this robustness in complex systems is a constrained and limited quantity and must be diligently managed.

The first explicit attempt to cast network design, modeling, and generation as a HOT problem was instigated by Fabrikant *et al.* [14]. They proposed a toy model of Internet growth in which each newly arriving generic node establishes connectivity by solving locally identical types of simple multi-objective optimization problems. A new node attempts to simultaneously optimize two objectives: “last mile” connection cost and “node centrality” cost. The “last mile” connection cost is supposed to capture the cost of resources associated with connecting to a parent node and is measured in terms of Euclidean distance. The “node centrality” cost reflects transmission delays and is measured in terms of the average hop distance from a potential parent to all other nodes in the graph. More formally, assuming a tree topology, when a node i joins a graph, it attaches itself to the node j that minimizes the weighted sum of the two objectives: $min_{j<i}(\alpha \cdot d_{ij} + h_j)$, where d_{ij} is the Euclidean distance between i and j, and h_j is the average hop distance from j to the entire graph. Depending on the relative importance of the two objectives in this multi-objective optimization, i.e., depending on the value of the parameter α, the authors prove in [14] that their HOT model yields three different regimes of graphs with qualitatively very different hierarchical structures and node degree distributions. Weighting the transmission delays over the connection cost (small values of α) creates star-like topologies, centered primarily around a single node (“hot spot”). If the connection cost dominates over the transmission delays (large values of α), the resulting graphs are more like random graphs, with exponential-tail type node degree distributions. The interesting regime lies in between (i.e., medium α-values) and consists of graphs with power-law type node degree distributions.

B. The Generic HOT Model and AS Connectivity

Although never stated explicitly in [14], the HOT model of Fabrikant *et al.* is generic in the sense that its nodes are neither routers nor ASs, and its links express neither router-level physical connectivity nor inter-AS peering relationships. Since our focus is on AS-level connectivity, we first ask whether the proposed HOT model is indeed appropriate and relevant to an AS-centric description of Internet growth.

When interpreting a graph’s nodes as ASs, the HOT model of Fabrikant *et al.* assumes a randomly-chosen *point* location for each AS and motivates the particular multi-objective optimization problem by arguing that in terms of connection cost, there exists a trade-off between being *attracted* to the topological core and being geographically *separated* from the core. However, a typical AS or ISP maintains in general multiple *Point of Presences* (PoPs) within its networks, and each PoP is a physical access location where customer ASs can connect to (e.g., see [18], [29]). Naturally, for a customer AS choosing among competing upstream providers, important considerations include: does the provider have a PoP nearby, and how many PoPs does the provider maintain globally [23]. Furthermore, large ASs close to the topological core can typically afford to invest more financial resources in their network infrastructure and by doing so, tend to further increase their global reach (i.e., number of PoPs and geographic diversity of PoP locations). Consequently, large ASs near



(a) Number of locations per node

(b) Node degree

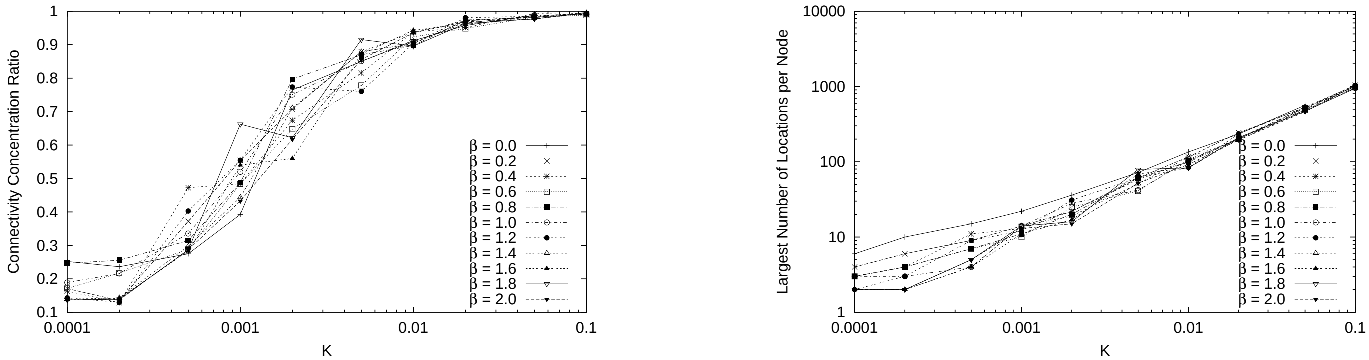
Fig. 3. The modified HOT model (α=5 and n=10,000)

the topological core are more likely to also be in closer geographic proximity to new customers than smaller ASs, which in addition are predominately located around the topological edges. This argument illustrates that when applying the HOT model of Fabrikant *et al.* directly to Internet growth at the AS level, the two proposed optimization objectives may no longer be *independent*—when nodes are multi-PoP ASs, a candidate node with low transmission delays has presumably also a low "last mile" connection cost. As a consequence, in a setting that is more realistic and allows for multi-PoP ASs, the HOT model of Fabrikant *et al.* can be expected to yield star-shaped graphs.

To test this hypothesis, we modify the generic HOT model of Fabrikant *et al.* to account for multi-PoP AS structures as follows. Each existing node u in the graph has a location list $loc\_list(u)$, which contains the set of its locations (as in [14], we work in the unit square, but as usual, the shape is inconsequential). For a new node i, $loc\_list(i)$ is initialized to contain a single randomly chosen location. To attach the new node i to the graph with i-1 existing nodes labeled 1 through i-1, instead of connecting it to target node j that satisfies the generic HOT criterion, i.e., $min_{j<i}(\alpha \cdot d_{ij} + h_j)$, we connect it to target node j that satisfies the modified HOT criterion: $min_{j<i}(\alpha \cdot (min_{l \in loc\_list(j)} d_{il}) + h_j)$. That is, when a newly added node searches for a node to attach to, it considers the same optimization problem as for the generic HOT model, except that each candidate node now has *multiple* locations. Thus, for each candidate node, its closest location has to be found first.

To reflect our intuition that with a growing customer base, ASs will increase their number of PoPs, and that the likelihood of a larger AS installing an additional PoP is greater than that of a smaller AS, after node i attaches to the graph, each existing node u (nodes 1 to $i-1$) is given a chance to increment the number of its PoPs by one. Specifically, with probability $p_{loc}(u)$, each existing node u adds a new random location to $loc\_list(u)$; with probability 1-$p_{loc}(u)$, no new location is added. The probability $p_{loc}(u)$ is given as $K \cdot rank(u)^{-\beta}$, where K and β are positive constants and $rank(u)$ is the rank of node u when all the existing nodes in the graph are sorted by the number of their children nodes in a monotonically decreasing order. The parameter K constrains the maximum number of locations per node and satisfies $0 \leq K \leq 1$; the maximum number of locations per node increases as K increases. On the other hand, β governs the decay of the distribution of the number of locations per node (i.e., exponential type decay for small β-values, power-law type decay for large β-values). **We call this the *modified HOT model* for Internet growth at the AS level.** In contrast to the generic model of Fabrikant *et al.*, this modified model allows for multi-PoP ASs and attempts to capture their evolution in time. As the connectivity of the graph grows over time, the internal structure (i.e., the number of PoPs and their geographic locations) of individual nodes can also expand.

Fig. 3 shows the results of three sample graphs generated by our modified HOT model (for three different (K, β) pairs; each graph contains 10,000 nodes). Plot (a) depicts the frequency plot of the number of locations per node, and the node degree frequency plot is shown in plot (b). In all three cases, the node with the highest degree (i.e., the highest ranked node in Fig. 3(b)) has acquired connections to a majority of all the other nodes (98.7%, 95.8%, 70.7% of all nodes, resp.), a clear indication that the resulting graphs exhibit a pronounced star-shaped structure. For comparison, we also plot in Fig. 3(b) the node degree distribution produced by the generic HOT model (i.e., a single location per node) and observe that in this case, the node with the highest degree is connected to only 14.5% of all nodes. This result confirms our hypothesis and illustrates that the optimization trade-off that is the basis of the generic HOT model can be seriously defeated in a more realistic AS networking setting.



(a) Connectivity concentration ratio (b) Largest number of locations per node

Fig. 4. Parameter (K,β) exploration in the modified HOT model (α=5 and n=10,000)

A more careful exploration of the parameter space (i.e., (K, β)-values) associated with the modified HOT model further confirms our conclusion. We define the *connectivity concentration ratio (CCR)* as the ratio of the largest node degree over the total number of links in the graph. Using CCR as a metric for assessing the degree of "star-shapedness" of a graph (e.g., CCR=1 for a perfectly star-shaped topology), we examined a set of graphs generated with a broad range of (K, β)-values. Fig. 4(a) plots the CCR value of the graphs generated by the modified HOT model with 0.0001 $\leq K \leq$ 0.1 and 0 $\leq \beta \leq$ 2.0. Fig. 4(b) shows the largest number of locations per node in the corresponding graphs. From Fig. 4(a), one can see that irrespective of the value of β, CCR converges quickly to 1 as K increases. In fact, graphs with CCR-values less than 0.5 or so are only possible for small K-values, which, according to Fig. 4(b), give rise to graphs where the largest number of node locations is only about 10 or less. Note that such limited variability in the numbers of node locations (small K values) renders the modified HOT model increasingly indistinguishable from the generic HOT model, where each node has only a single location.

IV. AS Geography as a Causal Force

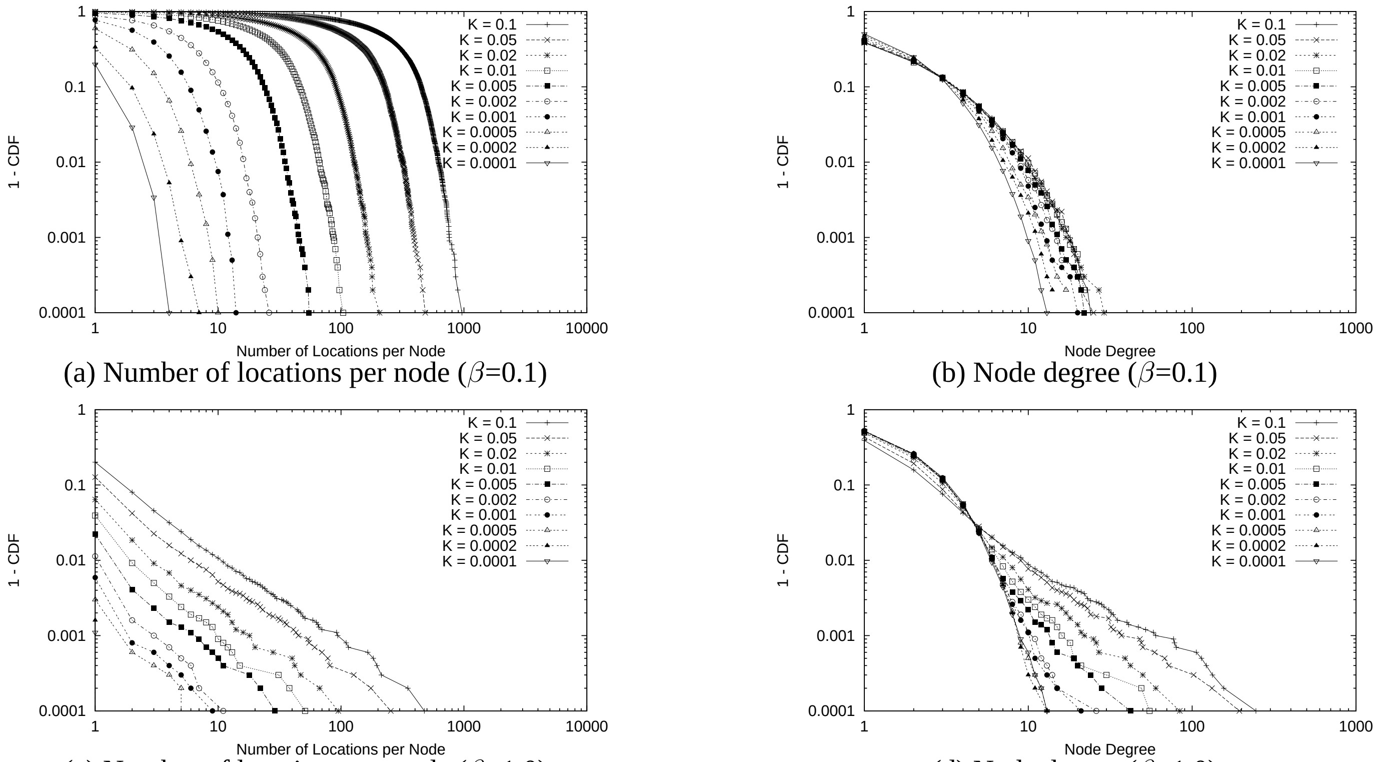
Motivated by networking reality (e.g., the need to allow for multi-PoP ASs with realistic geographies) and the shortcomings of the HOT model of Fabrikant *et al.* discussed in Section III, we describe in the following the first step of our construction of a new HOT model for Internet growth at the AS_{PC} level. To this end, we distill the role that AS geography plays in shaping AS connectivity, formulate a class of single-objective optimizations by which newly arriving customer ASs select their upstream provider, and validate the findings against relevant AS-specific data.

A. The Univariate HOT Model

To account for the multi-PoP structure of real-world ASs, we rely on the modified HOT model described in Section III-B, but redefine the local optimization problem by which an individual AS connects to the existing AS graph. In particular, viewing nodes as customer or provider ASs and links as provider-customer peering relationships and working as before in the context of a tree topology, we consider the single-objective optimization criterion consisting of simply minimizing the "last mile" connection cost. Under this growth model, each newly arriving node i is originally identified with a single PoP location; upon arrival, it always connects to the existing node j that contains the PoP location in $loc\_list(j)$ that minimizes the Euclidean distance to node i.[9] As in the case of the modified HOT model, each existing node u then gets a chance to enlarge its internal PoP structure by adding a new, randomly placed PoP, increasing thereby the geographic reach or diversity of u. For each node u, the probability to increase by one the number of its PoPs in $loc\_list(u)$ is again given by $p_{loc}(u)$. **We call this HOT model with its single locality-based connectivity objective the** ***univariate model*** **for Internet growth at the AS_{PC} level**.[10] It is ideally suited to test the hypothesis alluded to in Section III-B that the power-law type node degree distributions of inferred AS_{PC} or AS graphs may simply be due to the presence of power-law type distributions that capture in a parsimonious manner the high variability in the geographic extent of AS infrastructures. To recall, in Section III-B we argued that due to the proximity to the topological core of the AS graph and the geographic diversity of their PoP infrastructure, large ASs are more likely to acquire new ASs, which in turn enables them to build up their PoP in-

[9]Being an existing node, node j will in general already possess an extensive internal PoP structure.

[10]Univariate refers to the fact that geographic proximity is the only objective that is being optimized by a newly arriving AS.



(a) Number of locations per node (β=0.1)

(b) Node degree (β=0.1)

(c) Number of locations per node (β=1.0)

(d) Node degree (β=1.0)

Fig. 5. The univariate HOT model (n=10,000)

frastructure more aggressively than small ASs. Together, these arguments make it plausible that high variability in the number of PoPs per AS may well cause AS degrees themselves to exhibit high variability.

To test this hypothesis for a sensible range of the model parameters K and β associated with our univariate HOT model, we generate two sets of graphs, where each graph consists of 10,000 nodes. One set is produced by setting $\beta = 0.1$ and varying K between 0.0001 and 0.1, the other set is constructed with $\beta = 1.0$ and the same K range. The top row of Fig. 5 shows the distributions of the number of locations per node in plot (a) and the distributions of node degrees in plot (b) for the first set of graphs (i.e., $\beta = 0.1$). The bottom row of the figure shows the corresponding information in plots (c) and (d) for the second set of graphs (i.e., $\beta = 1.0$). Clearly, none of the distributions associated with the first set of graphs exhibit high variability or power-law type tails. In fact, all the distributions show extremely limited variability, fully consistent with exponential-type distributions. In particular, we observe that for graphs with $\beta = 0.1$, the exponential-type distributions for the number of locations per node are incapable of producing highly variable node degree distribution. However, graphs with $\beta = 1.0$ (bottom row) give rise to a more interesting behavior. As the value of K increases from 0.0001 to 0.1, not only does the number of locations per node become more highly variable, but the corresponding node degree distributions also change from exhibiting very limited variability (exponential-type) to showing high variability of the power-law type. In particular, for the larger K-values (e.g., $0.01 \leq K \leq 0.1$), there is a striking resemblance between the distributions of the number of locations per node and the node degrees. From this experiment, we conclude that our hypothesis has merit; that is, the geographic extent of existing nodes (as measured by the number of locations) can indeed be a controlling force in determining the characteristics of node degree distributions. This finding motivates a careful investigation into the geographic properties of existing ASs, with the ultimate goal of validating the high variability of AS size (measured in terms of the number of PoPs per AS) through elementary, geography-related measurements.

B. On Validating the Univariate HOT Model

The proposed univariate HOT model for Internet growth at the AS_{PC} level suggests that the high variability of the AS sizes as measured by the number of PoPs is an explanation for the striking power-law type AS degree distributions of inferred AS_{PC} graphs and, in turn, of the AS graphs themselves. To validate this aspect of the model and "close the loop" in the sense of [32], we provide below empirical evidence for the ubiquity of high variability

in the geographic extent of existing ASs. In particular, we present and discuss in the following the types of measurements and analysis methodologies that enable us to characterize various aspects of existing AS geographies, including the distribution of the number of PoPs per AS.

To infer AS size in terms of the number of PoPs, we first need to infer the geographic locations of the PoPs of the different ASs. To this end, we first collected a set of address prefixes from the Oregon route-views BGP data. Since BGP-advertised address prefixes can contain overlapping address space, to be able to associate address space with geographic location, we converted any overlapping address prefixes to a set of *disjoint* address blocks by recursively splitting in half any prefix that contains a subprefix. For example, given the overlapping address space 128.182.0.0/16 and 128.182.64.0/18, the recursive splitting procedure generates the three disjoint address blocks 128.182.128.0/17, 128.182.64.0/18, and 128.182.0.0/18. With 110,281 BGP-advertised address prefixes from the Oregon BGP data, after applying the recursive splitting, we were left with a total of 197,841 disjoint CIDR blocks.

Next, to associate a geographic location with each of these disjoint CIDR blocks, we randomly picked an IP address from each of the 197,841 blocks and queried the NetGeo database, a public repository of geographic information associated with address prefixes [10], for its geographic location. The NetGeo server responded with geographic location records for about 97% of all the queries we made. Relying on all these NetGeo records, we then map each of the geographically known IP address/CIDR block to its corresponding AS, thereby producing the geographic mapping information for the individual ASs.

While measuring the geographic extent of existing ASs by the number of their distinct NetGeo-inferred locations, we found that the geographic granularity of the NetGeo data is inconsistent, both with respect to $(longitude, latitude)$ coordinates and $(city/state/country)$ information.[11] To ensure that our inference methods are not biased or invalidated by these inconsistencies, we use a more flexible and practical definition of geographic granularity by relying on the following heuristic. For each AS, using the corresponding inferred geographic mapping information, we first obtain the $(longitude, latitude)$ coordinates of all of its distinct locations. Then we merge locations whose pairwise geographic distances are below a given threshold distance τ into one cluster or *location-group* and use the arithmetic mean of the $(longitude, latitude)$ coordinates of the individual locations that got merged as the cluster's new coordinates. Thus, by clustering closely linked geographic locations, our method is expected to recover, at least to a first order, the actual physical facilities of existing ASs. We will use in the following the number of inferred location-groups per AS as our estimate for AS size.

TABLE III

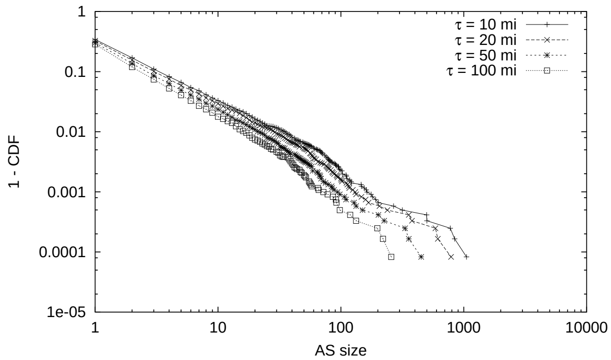
ROCKETFUEL VS. NETGEO

AS	Name	Number of PoPs	
		Rocketfuel	NetGeo (τ = 100 mi)
1221	Telstra	61	16
1239	Sprintlink	43	302
1755	Ebone	25	13
2914	Verio	121	51
3257	Tiscali	50	3
3356	Level3	52	42
3967	Exodus	23	36
4755	VSNL	10	35
6461	Abovenet	21	23
7018	AT&T	108	245

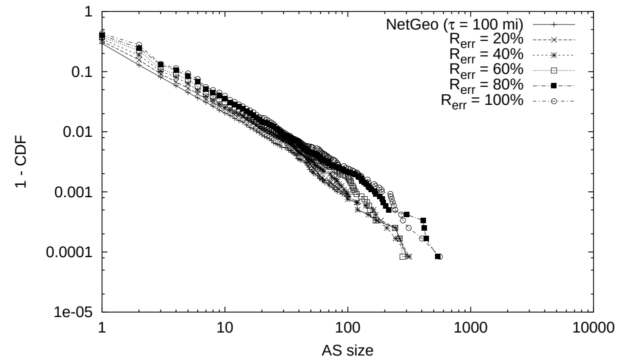
A number of recent studies have been concerned with inferring the geographic locations of existing network-related entities such as Internet hosts, IP routers, PoPs, ISPs, or ASs [26], [22], [29], [19], [10]. Of particular relevance is [29], where the authors, as part of obtaining some of the most complete currently available public router-level maps for 10 existing ASs, also inferred the number of PoPs for those same 10 ASs. To compare, Table III shows the inferred number of PoPs reported in [29] (left column) and obtained using our NetGeo-based approach (right column). Clearly, the differences are significant and beg an explanation. Consider, for example, an extreme case like AS 1239 (SprintLink), where [29] reports 43 PoPs while our method yields 302 (with τ = 100 mi). To understand why such differences are to be expected, first note that the geographic granularity of the two methods is different. While [29] uses city names, the geographic granularity of our method can be finer or coarser, depending on the choice of the threshold distance parameter τ. Second, a design feature of SprintLink is that small city customers tend to be back-hauled to far away PoPs (typically located in the major cities) through geographically close-by layer-2 switches which remain invisible to the approach pursued in [29], but which are likely to inflate the number of inferred PoPs when using our method (i.e., our method may wrongly infer the locations of such layer-2 switches as actual PoP locations).[12]

[11]For example, since NetGeo infers US locations from zip code or phone area code information (typically of finer granularity than city names), San Francisco has 40 distinct $(longitude, latitude)$ locations in the NetGeo data set, but they differ only slightly from one another. On the other hand, for most non-US cities, NetGeo assigned only a single geographic location.

[12]It could be argued that for the purposes of this paper, layer-2 switches in a SprintLink-like network design should be viewed and counted as PoPs.



(a)



(b)

Fig. 6. AS size distribution (size = number of PoPs).

One way to mitigate the inherently difficult problem of inferring geographic locality information for network elements (e.g., number of PoPs per AS) from BGP- or traceroute-derived measurements is to demonstrate that the obtained information is broadly robust. To illustrate, consider the inferred AS size obtained using our NetGeo-based technique. Plot (a) in Fig. 6 depicts the distributions of the inferred AS sizes for different values of the threshold distance parameter τ (i.e., 10, 20, 50, and 100 mi) and shows that AS size distribution as measured by the number of PoPs per AS is largely insensitive to the choice of τ. Plot (b) in Fig. 6 demonstrates an even stronger robustness property of the AS size distribution. To explicitly account for possible PoP count inflation (as in the case of SprintLink) or deflation (e.g., too coarse of a geographic granularity), we assume that R_{err}% of all the ASs have their sizes wrongly inferred by our NetGeo technique. We randomly mark half of them as having their sizes inflated, with the other half having deflated sizes. For each AS with an inflated size, we “adjust” its size by multiplying it by $1/\epsilon$, where ϵ is randomly chosen from $[1, \epsilon_{max}]$ and where ϵ_{max} is a parameter indicating some maximum margin of error. On the other hand, for each deflated AS, we “modify” its size by multiplying it by a factor of ϵ, where ϵ is again picked at random from $[1, \epsilon_{max}]$. For example, if the inferred size of AS X is 100 and $\epsilon_{max} = 4$, the adjusted size of AS X can vary between 25 and 100 if X is considered inflated, and between 100 and 400 if it is marked deflated. Assuming that the inferred AS size can deviate from the (unknown) “true” size by as much as a factor of 4 (i.e., $\epsilon_{max} = 4$), plot (b) in Fig. 6 testifies to the robustness of the power-law type behavior of the AS size distribution with respect to a wide range of misspecifications of AS sizes (i.e., R_{err} = 20%, 40%, 60%, 80%, and even for R_{err} = 100%). Note that such robustness properties indicate an enormous resilience to ambiguity of the unknown underlying distribution (e.g., ambiguity with respect to precisely how AS size is defined or measured) and essentially imply that the latter must be of the power-law type [9]. Based on these empirical findings, we conclude that the power-law type distribution for AS sizes that has been identified by our univariate HOT model for Internet growth at the AS_{PC} level as an explanation for the power-law type node degree distributions of inferred AS_{PC} and AS graphs is fully consistent with data derived from geographic mapping information for the individual ASs.

C. Some comments on using the NetGeo database

[22] reports on a study of the geographic population of Internet routers that were discovered by traceroute probing. That work relies on proprietary geographic location databases to infer the coordinates of all router interface addresses. The work described here differs from [22] in a number of ways. For one, as discussed in Section 4.2 above, instead of considering an inevitably incomplete set of router addresses obtained from traceroute probing, we examine the locations associates with all existing address prefixes. Second, we rely on a public repository of geographic information for IP address prefixes, i.e., the NetGeo database [10].

A typical NetGeo record is shown below and contains $(city/state/country)$ information and the $(latitude/longitude)$ coordinates for the (target) IP address in question.

```
LONG: -77.11
LAT: 38.87
STATE: VIRGINIA
NAME: CHY-WAN-65-121-98A
STATUS: OK
COUNTRY: US
LAT_LONG_GRAN: City
TARGET: 65.121.98.28
NUMBER: 65.121.98.0 - 65.121.98.255
CITY: ARLINGTON
NIC: ARIN
LAST_UPDATED: 14-Nov-2001
LOOKUP_TYPE: Block Allocation
DOMAIN_GUESS: qwestip.net
```

The geographic location records of the NetGeo database originate from the ARIN/APNIC/RIPE whois servers. The NetGeo database server extracts geographic information such as city, state, country names from the text of whois records. It also leverages US zip codes, phone area codes, or email addresses with 2-letter top-level domains (TLDs) country code where possible. For large, geographically-dispersed transit providers for which the whois-based technique is not sufficient for location mapping, the NetGeo obtains geographical hints from DNS lookup as well.

Given that NetGeo's inference engine exploits various geographic cues, we performed a careful analysis to calibrate its mapping accuracy. The following are the main conclusions of our assessment of the reliability of the NetGeo data.

- *Comparison with Geotrack (US sites only)*: Using 1,000 US university sites with known geographic locations, comparing their NetGeo-inferred locations with the corresponding locations inferred by the Geotrack method (see for example [26]), we found that the NetGeo-based locations are slightly more accurate than those produced by Geotrack (while the median difference between the former and the actual locations was 0, it was about 80 miles for the Geotrack-related differences).
- *Comparison with Geotrack (random sites)*: Using 660 randomly chosen IP addresses whose actual geographic locations are typically not known, we observed that the NetGeo- and Geotrack-inferred locations vary widely, with as many as 40% of the IP addresses having NetGeo- and Geotrack-inferred locations that are more than 1,000 miles apart from one another. It turns out that the reason for this extreme discrepancy between the two inference methods is that Geotrack has a tendency to associate the geographic location of networks that reside outside the US with the location of US cities where the the networks' international circuits originate or terminate.
- Comparison with Geotrack (random US sites): When considering only those random IP addresses (out of the above 660) which – according to NetGeo – are located within the US, we found that the NetGeo- and Geotrack-inferred locations are now more comparable, with a median difference of about 100 miles.

As described in Section 4.2, because of inconsistencies in the geographic granularity of the NetGeo data, we consider a flexible and very practical definition of geographic granularity that essentially ignores location information that is too fine-grained (i.e., below a threshold distance of τ miles). As a by-product, this definition can also mitigate the problem whereby portions of address space belonging to an ISP X may be delegated to other business customers that obtain their network access from X without running

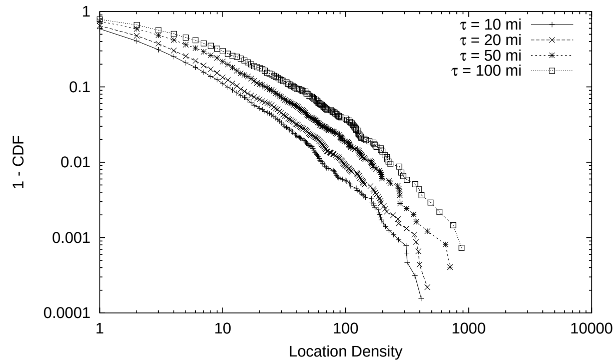


Fig. 7. Location density distribution (density = number of ASs).

BGP themselves (e.g., managed hosting service clients or small dial-up/broadband access sites). In this case, the NetGeo data set may associate the locations of such re-assigned address blocks with those of these business customers, causing the location of ISP X to be wrongly extended to the locations of those customer sites. We predict that such business clients would typically be located in geographic proximity to their provider ISP's actual locations due to the economic incentive of laying shorter lines or the ease of administrative contact. In any case, the heuristic is likely to track the locations of the physical facilities of a given AS with reasonable accuracy.

We conclude this section with another example of how the NetGeo data can be used to study interesting features associated with AS geography. Recall that the development of our locality-basd HOT model for Internet growth at the AS_{PC} level was partly motivated by the argument that large multi-PoP ASs are more likely to acquire new customers than single-PoP or small multi-PoP ASs. Knowing the geographic concentration of existing ASs (i.e., the number of different ASs per geographic location) would shed light on the validity of the above argument. Thus, to obtain the number of ASs per geographic location, we first collect from our NetGeo data set the ($longitude$, $latitude$) coordinates of all the distinct geographic locations that are associated with at least any one AS. As before, depending on the threshold distance τ, we group all the pairwise close geographic locations into one cluster or location-group and then associate individual ASs with the different location-group as follows: AS i is associated with a given location-group if and only if AS i has at least one geographic location that is contained in the location group. When equating the number of ASs per location with the inferred number of ASs per location-group, Fig. 7 shows the resulting distributions for four different values of the threshold distance τ (i.e., τ =10, 20, 50, and 100 miles). Additional details about the top-10 locations in terms of the number of ASs

TABLE IV
TOP-10 LOCATIONS (τ = 10 MILES)

Location Rank	Location Name	Number of ASs
1	New York City, NY, US	508
2	Herndon, VA, US	411
3	Amsterdam, NL	366
4	Fairfax, VA, US	316
5	Englewood, CO, US	313
6	Seoul, KR	310
7	Palo Alto, CA, US	269
8	Atlanta, GA, US	245
9	Los Angeles, CA	225
10	London, UK	211

present can be found in Table IV. Not surprisingly, the top-10 locations correspond mainly to large cities or major urban areas, with high population densities and significant economic activities.[13] From Fig. 7 we can conclude that the geographic concentration of existing ASs exhibits high variability (ranging over 3 orders of magnitude) and is largely insensitive to the details of defining geographic granularity (i.e., specific choice of τ).

V. AS BUSINESS MODEL AS A CAUSAL FORCE

Building upon our univariate HOT model, we describe in the section the second step of or construction of a new HOT model for Internet growth at the AS_{PC} level. In particular, we explore here in detail the role that the decision making processes by which the different customer ASs enter into business relationships with their provider ASs play in shaping Internet AS connectivity.

A. Inter-AS Peering in the Commercial Internet

By our univariate HOT model, the variability in AS size roughly determines the variability in AS node degrees (see Fig. 5). However, while inferred AS graphs of the actual Internet (e.g., Fig. 1) typically show 3 to 4 orders of magnitude variability in node degrees, the variability of the inferred number of PoPs per AS (i.e., AS sizes) tends to range over only 2 to 3 orders of magnitude (see Fig. 6). While this suggests that managing ASs with too many PoPs is economically or technically not a viable business, it is also a clear indication that our univariate HOT model has only identified a necessary but by no means sufficient condition for explaining the high variability of inferred AS node degrees. The model's claim is fully consistent with elementary, AS geographic measurements, but it does not rule out the existence of factors other than geographic proximity at work in shaping AS connectivity.

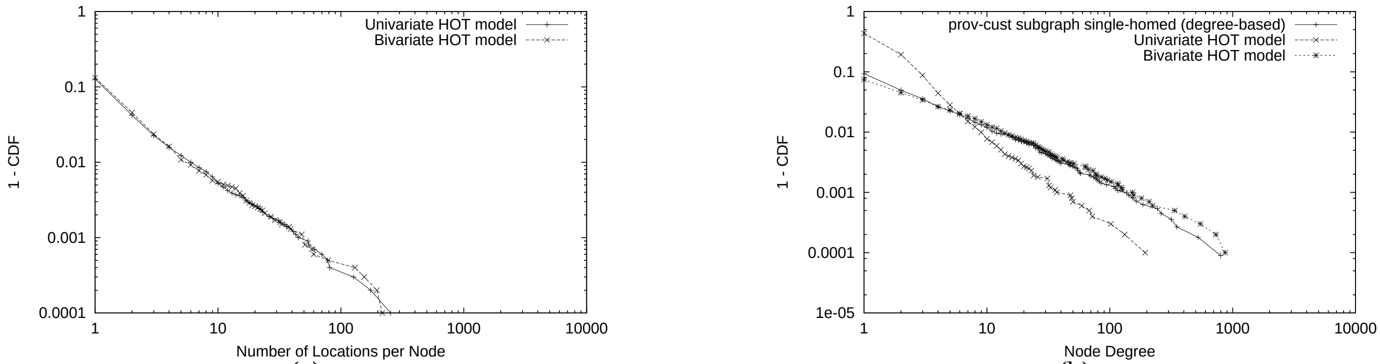
One such factor is how a new customer AS X establishes Internet connectivity when there are numerous provider ASs within the same (or comparable) geographic proximity of X, all vying to become AS X's provider.[14] In such situations, it is reasonable to assume that AS X will evaluate the competing providers by several criteria other than geographic proximity (including for example the availability, reliability, and performance of the networks, available pricing plans, existing customer support, number of value-added services, geographic reach and projected network build-out, prior acquaintance between the parties involved, etc. [23]) and then make a more or less rational decision to choose a provider that is "best" or "optimal" with respect to some, possibly AS-dependent, utility measure. To formalize this admittedly over-simplified process by which new customer ASs select their upstream provider, we view the above criteria as being part of an abstract object called AS X's *business model*. Aside from the above mentioned criteria, the business model of an AS can include any other factors that may play a role in how AS connectivity is established. In Section V-B below, we present a concrete mathematical formulation of such toy business models, flexible enough to account for a range of different business objectives among the various ASs.

B. The Bivariate HOT Model

In incorporating an AS's *business model* into our proposed HOT model, we assume that a newly arriving AS first identifies within its geographic neighborhood all provider ASs offering service. The candidate providers are then evaluated based on a set of criteria in the AS's business model. Finally, the "best" upstream provider for the new AS is selected as the result of the AS's locally optimal business decision. More formally, we start with the univariate model defined in Section IV-A, where each node i has a list $loc\_list(i)$ that provides geographic information about its internal PoP structure (the number of PoPs and their geographic coordinates). To define node i's business model, we augment node i with an N-dimensional vector $\boldsymbol{x}_i = (x_{i,1}, x_{i,2}, ..., x_{i,N})$ and a non-empty set $S_i \subset \{1,2,...,N\}$. We call $\boldsymbol{x}_i$ node i's *score vector*. An instantiation of $\boldsymbol{x}_i$ is obtained by assigning each of its components $x_{i,n}$ a uniform random number taken from [0,1]. S_i is node i's *selection set* and and is randomly chosen from $2^{\{1,2,...,N\}} \setminus \{\phi\}$. The pair $(\boldsymbol{x}_i, S_i)$ forms node i's *business model*. Intuitively, the score vector indices represent the possible criteria that enter into the decision making processes by which competing ASs are *evaluated against*. The individual vector-components $x_{i,j}$ quantify how AS i

[13]As we vary τ from 10 miles to 100 miles, we notice a slight reordering among the top-10 ranking, but the ASs remain the same.

[14]The importance of this factor in establishing provider-customer peering relationships among ASs is emphasized by the observed highly variable AS density per geographic location as discussed at the end of Section IV-B.



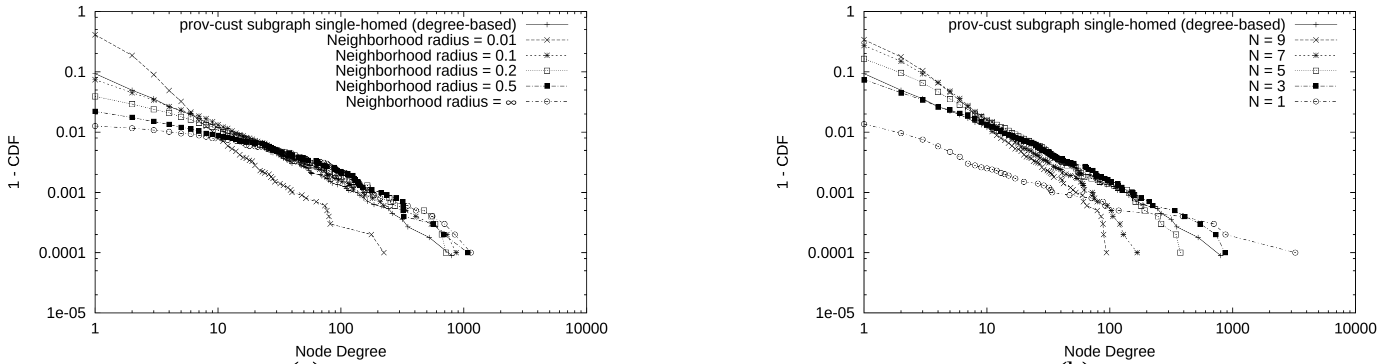
(a) (b)

Fig. 8. The bivariate HOT model (n=10,000, K=0.05, β=1.0, neighborhood radius=0.1, N=3)

is measured up with respect to criterion j, $1 \leq j \leq N$. The selection set S_i defines the subset of criteria that AS i deems relevant when choosing its upstream provider. Hence in making its choice for a provider, a new AS i matches up its selection set against the score vector of each candidate provider. The selection set is generally different from one AS to another.

To determine the comparative desirability of provider ASs, we define *node dominance on a set P* as follows. Given two nodes i and j and a non-empty set $P \subset \{1, 2, ..., N\}$, node i is said to *dominate* node j on set P if $x_{i,n} \geq x_{j,n}$, for all $n \in P$, and $x_{i,m} > x_{j,m}$, for some $m \in P$. To illustrate, consider the following simple example whereby three competing provider ASs, AS X, Y, and Z, are evaluated by a newly arriving customer AS i in terms of two ($N = 2$) criteria: network reliability and unit bandwidth cost. AS X reportedly has 99% network reliability, AS Y 98%, and AS Z 97%. They charge \$100/mon., \$150/mon. and \$50/mon. per unit bandwidth, respectively. If the new customer AS i considers unit bandwidth cost as the only relevant criterion in choosing a provider (i.e., S_i = {unit bandwidth cost}), then it will choose AS Z. However, a new customer AS j that deems both network reliability and unit bandwidth cost important when selecting its upstream provider (i.e., S_j = {network reliability, unit bandwidth cost}) will *not* select AS Y, since AS Y is dominated by AS X. It will choose either AS X or Z, for neither is dominated by the other (ties can be broken as j pleases). Note that in this process, the score vector of a newly arriving customer itself is not needed, but its selection set is crucial; at the same time, any newly arriving customer AS needs to know the score vector $\boldsymbol{x}_u$ and geography data $loc\_list(u)$ of each existing provider AS u in the graph.

According to this model, the graph grows as follows. When a new node i arrives, it is assigned the trivial list $loc\_list(i)$ (consisting of i's coordinates only) and an instantiation of its selection set S_i. Node i first initializes its candidate set $candidate(i)$ with all existing nodes that have a PoP location within a pre-defined Euclidean distance from node i, the so-called *neighborhood* of i. Next, for each pair of nodes in $candidate(i)$, if one of the nodes is dominated on the set S_i by the other node, the new node i marks that dominated node. Subsequently, all marked nodes are eliminated from $candidate(i)$. The new node then randomly picks one of the nodes left in $candidate(i)$ as the target node and establishes a connection to that node. Finally, node i is given an instantiation of its score vector $\boldsymbol{x}_i$, and—as in the multi-PoP univariate HOT model—each existing node u is given a chance to increments the number of its PoP locations in $loc\_list(u)$ by 1, with probability $p_{loc}(u)$; with probability $1 - p_{loc}(u)$, $loc\_list(u)$ is left unchanged. **We call this HOT model in which two objectives (i.e., geographic proximity and economic utility) are optimized simultaneously the *bivariate model* for Internet growth at the AS_{PC} level**. This model differs from the previously considered univariate model in that a new AS i, instead of finding the geographically closest AS, now considers *all* the ASs with PoP locations "close-by," i.e., located within its neighborhood. More importantly, when trying to narrow down its choices among multiple locally accessible candidate provider ASs, the new AS i avoids selecting a provider that is dominated by some other competing provider in terms of the criteria present in its selection set S_i. The model ensures that the final selection of upstream provider by each new AS i is *Pareto optimal* [16] in the sense that no other "close-by" provider AS dominates the chosen provider on the set S_i. This HOT formulation is a truly multi-objective optimization, combining geography- and economic-specific objectives. Moreover, it defines a fully distributed and decentralized design of Internet connectivity at the AS_{PC} level, typically with different optimization problems (as a result of the different AS-specific selection sets) solved by the different ASs.



(a) (b)

Fig. 9. The bivariate HOT model (n=10,000, K=0.05, β=1.0) – (a) the effect of neighborhood radius (for fixed N=3), (b) the effect of the dimension of the score vector (for fixed neighborhood radius=0.1).

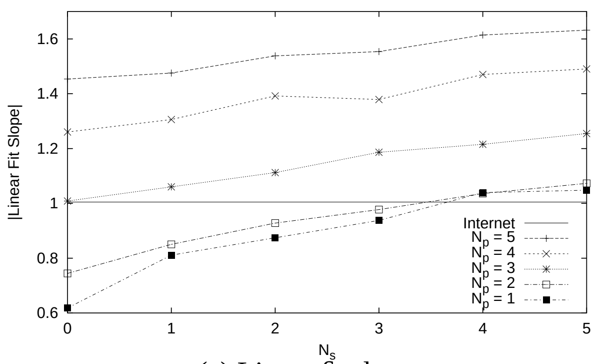
To demonstrate the potential of this bivariate HOT model, Fig. 8 illustrates some of the features of the resulting graphs (as always, each graph consist of 10,000 nodes). In particular, defining the neighborhood of a given node to be a disk of radius 0.1 centered at a node, setting $K = 0.05$, β=1.0, and $N = 3$, plot (a) of Fig. 8 depicts the distribution of the number of locations per node, and plot (b) gives the node degree distribution of the generated graphs. On the same plots, we also show the results corresponding to the graph generated by the univariate model (same parameters: $K = 0.05$, β=1.0, 10,000 nodes). Finally, for comparison, plot (b) also contains the node degree distribution corresponding to one of the single-homed provider-customer subgraph of the OREGON+ AS graph (i.e., the largest tree topology rooted at AS701; see Section II-C). Fig. 8 provides convincing evidence that while the univariate and bivariate HOT models yield comparable AS size distributions, the latter has a clear impact on the distribution of node degrees. In fact, it can increase the variability of node degrees by at least one order of magnitude beyond what the univariate model is capable of, closely matching the node degree distribution of inferred AS (sub)graphs.

Next we illustrate that both objectives in the formulation of our bivariate HOT model (i.e., geographic proximity and locally Pareto-optimal business decision) are necessary for creating graphs with Internet-like features. We first examine in plot (a) of Fig. 9 how the geography-related factor, i.e., the neighborhood of a node, impacts the node degree distribution of the resulting final graph. To this end, we set K=0.05, β=1.0, and N=3, but vary the radius of a node's neighborhood from 0.01, 0.1, 0.2, 0.5, and some large number. As can be observed, as the neighborhood size decreases to a disk of radius 0.01, the resulting node degree distributions exhibit less variability and become comparable to those generated by similarly parameterized versions of our univariate models (see for example, Fig. 8(a)). On the other hand, too large of a neighborhood size (e.g., neighborhood sizes ≥ 0.2) results in degree distributions that are qualitatively different from those derived from the actual Internet. These observations suggest that a certain degree of geographic locality is needed to model Internet-like AS_{PC} graphs with our bivariate HOT model.

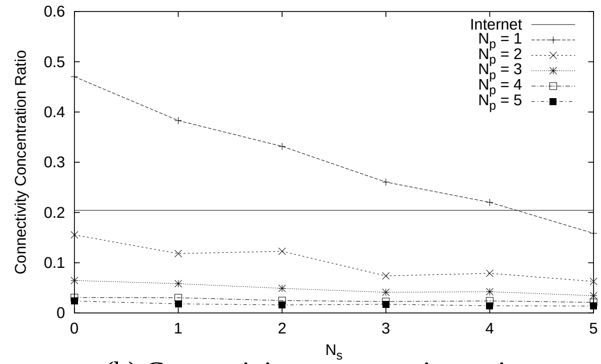
To examine how the business-related criterion affects node degree distributions, we parameterize our model with $K = 0.05$, β=1.0, and a neighborhood radius of 0.1, but vary the parameter N, the dimension of the score vector, from 1 to 9. N captures the complexity of the business-related aspect of provider selection and is therefore a natural parameter to consider for the purposes of this experiment. Plot (b) of Fig. 9 shows the node degree distributions of a 10,000-node graph, for the different values of N. As the value of N increases, the variability of node degrees tends to decrease. This observation agrees with the intuition that a small N-value creates more opportunities for some nodes to be favored over the others during the growth process of the graph, which in turn increases the variability of node degrees. Small N-values reflect the regime whereby relatively simple business decisions are involved in selecting the "best" provider. If complex business decisions are allowed (large N-values), the cardinality of the Pareto-optimal candidate set is likely to be large, causing individual node degrees to be more evenly distributed among existing nodes, thereby reducing the overall node degree variability.

C. On the Robustness of the Bivariate HOT Model

Our bivariate HOT model for Internet growth at the AS_{PC} level attempts to capture the local decision processes performed by individual ASs within an abstract but intuitive framework. In accordance to the model validation framework advocated in [32], the natural next step would be to empirically validate the model using rele-



(a) Linear fit slope



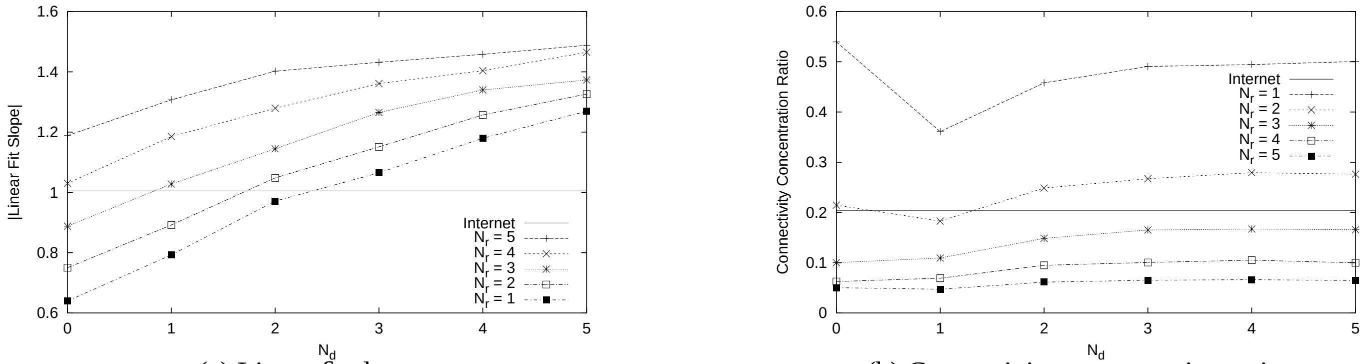
(b) Connectivity concentration ratio

Fig. 10. Primary vs. secondary criteria

vant data. However, at this point it befuddles us how to collect the appropriate data; in fact, we are not even sure what sort of data might be available that could shed some light on the validity of the proposed model, especially as far as the assumed toy business model is concerned. Fully realizing this dilemma, the next-best approach in support of the overall viability of the model is to assess its sensitivity to changes in the underlying toy business model $(\boldsymbol{x}_i, S_i)$ associated with AS i. The exploration of the impact of such changes or of related what-if scenarios is typically motivated by real-world inter-AS peering considerations and inevitably complicates the resulting business model beyond our intuitive but naive toy model, where—in the absence of empirical evidence to the contrary—the assignments of score vectors and selection sets are completely left to chance (i.e., independent and uniformly random). For the viability of our proposed bivariate model, it is therefore important to understand what sort of changes, refinements, or modifications to our toy business model result in qualitatively the same graph structures, and which disturbances give rise to qualitatively very different topologies.

Primary vs. secondary criteria To illustrate, the following refinement of the bivariate HOT model is motivated by the observation that when evaluating different provider ASs, customer AS i may be more concerned with some criteria (e.g., unit bandwidth cost) than others (e.g., customer support). This then suggests considering selection sets S_i for AS i that are not necessarily randomly chosen from $2^{\{1,2,\ldots,N\}} \setminus \{\phi\}$. To achieve this goal, we associate with a given node i an N-tuple of weights $\Delta(i)$ = $(\delta_{i,1}, \delta_{i,2}, ..., \delta_{i,N})$, $0 \le \delta_{i,n} \le 1$, where $\delta_{i,n}$ is the weight with which the node i indicates its preference to have criterion n included in its selection set. We then define the probability that the node's selection set S_i is instantiated with a particular subset X in $2^{\{1,2,\ldots,N\}} \setminus \{\phi\}$ to be $((\prod_{u \in X} \delta_{i,u})(\prod_{v \notin X} (1-\delta_{i,v})))/(1-\prod_{w=1}^{N}(1-\delta_{i,w}))$. By varying the weight-vector $\Delta(i)$, we are able to explicitly control the relative preference for including certain criteria (or indices of the score vector) in a selection set over others. For example, setting all weights equal to 0.5 corresponds to the purely random choice of selection sets assumed in our original formulation of the bivariate model. On the other hand, having $\delta_{i,n}$ larger than $\delta_{i,m}$ gives rise to selection sets that are more likely to contain criterion n than m. Given a weight-vector $\Delta(i)$, we next classify the score vector indices or criteria (but not the actual values of the components of the score vector) into *primary* and *secondary* indices, depending on whether $\delta_{i,n} \ge 0.5$ or $\delta_{i,n} < 0.5$, resp. That is, a primary index represents a criteria that an AS deems important when choosing its upstream provider; secondary indices correspond to criteria that the AS considers to be relatively unimportant.

In our experiment, we tested a special case where $\Delta(i)$ = $\Delta(j)$ = Δ for every pair of node i and j (i.e., the same weight-vector Δ for every node), and all primary indices are assigned equally high weights (i.e., 0.8) and all secondary indices equally low weights (i.e., 0.2). The number of primary and secondary indices in Δ is then denoted by N_p and N_s respectively, where N_p + N_s = N. Under this setting, we generated a set of graphs varying N_p from 1 to 5 and N_s from 0 to 5. Fig. 10(a) shows the inferred α-values (i.e., estimated tail-index of power-law distributions of the form $P[X > x] \sim x^{-\alpha}$, for large x) of the degree distributions of these generated graphs, where the α-estimates are the slopes of least squares fits on log-log scale). Fig. 10(b) plots their connectivity concentration ratios (CCRs) introduced in Section III-B. The reported values in both figures are averages taken over ten realizations that were generated for each pair of (N_p, N_s)-values. For comparison, we added to the figures the inferred α- and CCR-values of the P1 subgraph of OREGON AS graph (labeled "Internet").



(a) Linear fit slope

(b) Connectivity concentration ratio

Fig. 11. Regenerative vs. degenerative criteria

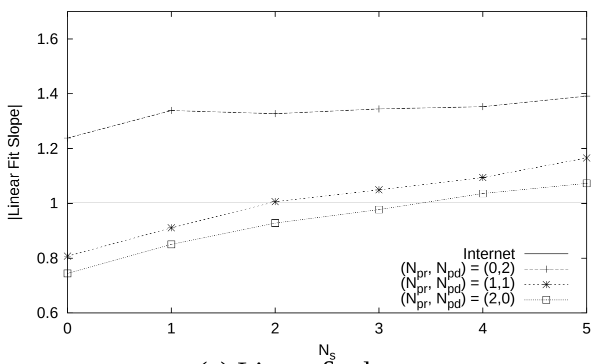
Fig. 10(a) suggests that as N_p or N_s increases, the slope of the degree distribution becomes steeper.[15] The reason for this behavior is similar to what explains Fig. 9(b). Due to different selection preferences for primary and secondary criteria, however, an increase of N_s has much less of an effect on the degree distribution than an increase of N_p. In addition, Fig. 10(b) shows that if $N_p > 1$, the CCR of a generated graphs falls short of that of the Internet AS graph. From this experiment, we conclude that when allowing for a separation of the score vector into components representing primary and secondary criteria, Internet-like node degree distributions are typically achieved with higher-dimensional score vectors than in our original model. Furthermore, to retain an Internet-like hierarchical structure (as measured in terms of the connection concentration ratio CCR), the value of N_p is required to be *smaller* than the dimension of the score vectors in our original model.

Regenerative vs. degenerative criteria Another type of refinement to our toy business model concerns the actual values assigned to the ASs' score vectors. This refinement reflects an aspect of networking reality whereby for example network reliability of a customer AS cannot be higher than that of its upstream providers', mainly because network access failures inevitably propagate through downstream customer networks. We call criteria whose quality for the customer AS is provider dependent *degenerative* criteria. In contrast, an AS's quality of customer support is a criterion that is not necessarily determined by its upstream providers'; we call such provider-independent criteria *regenerative* criteria. The introduction of degenerative criteria leads us to consider the impact of score vectors that are not necessarily independent across different ASs and whose components are not necessarily chosen uniformly between 0 and 1. In our original bivariate model, the score vector $\boldsymbol{x}_i$ of node i is determined independently of that of its parent node j's. If criterion m happens to be degenerative, then the value of the score vector component $x_{i,m}$ is chosen uniformly from $[0, x_{j,m}]$ (i.e., it depends on the particular AS j chosen as i's provider and on j's score vector). Regenerative criteria are assigned random values in $[0, 1]$ as in the original bivariate model, independent of everything else. The number of degenerative and regenerative criteria are denoted by N_d and N_r respectively, where $N_d + N_r = N$.

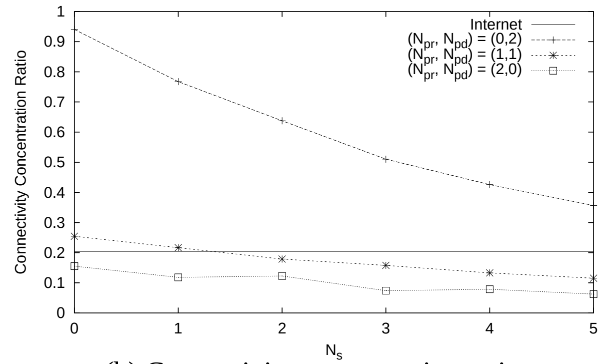
To explore the parameter space associated with this second refinement of our original bivariate model, we vary N_r from 1 to 5 and N_d from 0 to 5 and consider the case of uniformly random selection sets. Fig. 11(a) is consistent with our earlier observations that higher dimensional score vectors result in steeper degree distributions. In addition, one can see that increasing N_d while fixing N does not change the slope of the degree distribution significantly. For example, the four dots corresponding to $N = 4$ remain close to the horizontal reference line. Fig. 11(b) indicates that a higher N_d yields a graph with a higher CCR value. The effect of changing N_d is more dramatic when N is fixed. For example, incrementing N_d from 0 to 3 with $N = 4$ causes the CCR of a generated graph to increase from 0.05 up to 0.5. With more degenerative criteria, customer ASs are more likely connect directly to nodes closer to the root, resulting in an overall topology that becomes more star-shaped from the root node outward. The main finding from this set of experiments is that the presence of degenerative qualities or criteria provides a direct "dial" for impacting the growth dynamics of the graphs' cores as measured by the CCR metric.

Two refinements in combination Finally, we also experimented with combining the previous two refinements and allowed ASs to classify their criteria to be primary or secondary and, at the same time, degenerative or regenerative. Out of N_p primary criteria, N_{pd} criteria have de-

[15] In all cases reported in the figures, the quality of the linear fit is very high with correlation coefficient > 0.98.



(a) Linear fit slope



(b) Connectivity concentration ratio

Fig. 12. Primary regenerative vs. primary degenerative criteria

generative property and the rest of N_{pr} criteria have regenerative property (N_{pd} + N_{pr} = N_p). Likewise with N_s secondary criteria (N_s = N_{sd} + N_{sr}). Fig. 12 shows the α-estimates of the degree distributions and the CCR-values of graphs generated with N_p = 2 and 0 $\leq$ N_s $\leq$ 5. Primary criteria are provided with either degenerative or regenerative property, but all secondary criteria are classified as regenerative by default (i.e., N_s = N_{sr}).[16] Fig. 12(b) shows that if all primary elements have degenerative property (i.e., (N_{pr},N_{pd})=(0,2)) and N_s is small, the resulting graphs tend to favor a pronounced star-shaped topology. On the other hand, if all primary criteria have the regenerative property (i.e.,(N_{pr},N_{pd})=(2,0)), the CCR of a generated graphs fall short of that of the Internet AS graph. In short, neither of the two extreme cases yields the Internet-like hierarchical structure expressed by CCR. As long as we avoid such extreme cases, the refined model results in the Internet-like node degree distribution and hierarchical structure (e.g., (N_{pr},N_{pd})=(1,1) and N_s = 2). We made similar observations for $N_p > 2$ as well.

The main conclusions from experimenting with these refinements, either individually or in combination, are that the qualitative aspects of the resulting graphs (as captured for example by the node degree distribution or the connectivity concentration ratio CCR introduced in Section III-B) by and large agree with those obtained using the original bivariate model. Quantitatively, the graphs differ in very intuitive and predictable ways as a result of fine-tuning their growth dynamics through particular choices of the parameters associated with each refinement; that is, the number of primary and secondary, or degenerative and regenerative criteria. Note that these parameters provide explicit "knobs" for manipulating the sizing of the AS-specific sets of Pareto-optimal candidate nodes, which in turn determines node degree variability and hierarchical structure of the resulting graphs. For example, when allowing for degenerative and regenerative criteria, with more degenerative criteria, customer ASs are more likely to connect directly to nodes closer to the root, resulting in an overall topology that becomes more star-shaped, with the root at the center. On the other hand, when ASs are given the choice to classify their criteria to be regenerative or degenerative and, at the same time, primary or secondary, we find that when customer ASs select among a number of competing provider ASs, neither the primary regenerative nor the primary degenerative criteria of the providers dominate the local decision processes of customer ASs, but instead, *both* types of criteria are actively considered in their decision-making processes.

[16] We also tested the cases where some secondary criteria have degenerative property. However, as far as secondary criteria are concerned, whether they are degenerative or regenerative does not make much difference.

VI. AS Graph Evolution as a Causal Force

The third and final step of our construction of a new HOT model for Internet growth at the AS_{PC} level concerns primarily the graph growth process itself. This step is motivated by the observation that while the bivariate (as well as the univariate) model is purely incremental in nature (i.e., nodes are added to the graph one by one, and once added, they stay forever, and so do all the added links), the historical evolution of AS-related connectivity is obviously more dynamic, given the business dynamics of the ISP market.

A. The Multivariate HOT Model

In incrementally grown graphs like the ones generated by our bivariate HOT model, the parent (or ancestor) node of any given node i is always added to the graph before node i. In the Internet, however, existing ASs can and do disappear as a result of, for example, bankruptcies, often times leaving their customers scrambling for new provider ASs. To account for such change-of-provider scenarios

in our model, we consider the bivariate HOT model defined in Section V-B and introduce *node death events* that trigger a particular *change-of-provider* mechanism. More precisely, every time P_d new nodes have been added to the graph, one of the existing nodes u in the graph is randomly selected and removed from the graph (together with all its incident links). In turn, all of node u's children (if any) become orphans and have to select a new parent node to re-establish graph connectivity. When determining its new provider, the local decision process of a given orphan node i is similar to the one by which a newly arriving node connects to the graph in our bivariate HOT model, but differs in two important ways. First, when an orphan node i chooses its set of "close-by" candidate providers, none of the potential new parent nodes can be descendants of node i. Second, the definition of geographic proximity needs to be modified because in contrast to a newly arriving node u that is assigned a $loc\_list(u)$ containing a single PoP location, the orphan node i's internal PoP structure as described in $loc\_list(i)$ may have grown substantially since node i's birth. To account for this latter complication, we define the geographic proximity between an orphan node i and a potential new parent node j to be $(\sum_{k \in loc\_list(i)}(min_{l \in loc\_list(j)} d_{kl}))/|loc\_list(i)|$, the *expected* minimum distance between nodes i and j. If the orphan node i happens to have only a single location, this expression reduces to $min_{l \in loc\_list(j)} d_{il}$, the original definition of geographic proximity used in our univariate/bivariate model. **We call this final HOT model with its multi-faceted objectives (i.e., AS geography, AS business model, and AS evolution) the *multivariate model* for Internet growth at the AS_{PC} level.**

To judge the impact of allowing for evolutionary changes through the introduction of node death events and its associated change-of-provider mechanism, we generated two graphs (each with 10,000 nodes).[17] The first graph was generated using our bivariate HOT model (with $K = 0.05$, $\beta = 1.0$, and $N = 3$); the second graph resulted from an identically parameterized version of our multivariate HOT model, where in addition, we set $P_d = 5$.[18] While the node degree distributions of the resulting graphs are practically identical (not shown), the scatter-plots of node age vs. node degree depicted in Fig. 13 illustrate the impact that the introduction of change-of-provider events has on the graph growth process itself. We observe that by allowing for these specified evolutionary changes, the multivariate model in plot (b) introduces, as expected, more randomness between node ages and node degrees than the bivariate model in plot (a), thereby reducing the correlation between node age and degree to be more consistent with that observed in inferred AS_{PC} graphs. Note that in the case of these generated graphs, node age was calculated based on an empirically derived AS birth rate (see appendix for details).

B. On the Historical AS Evolution

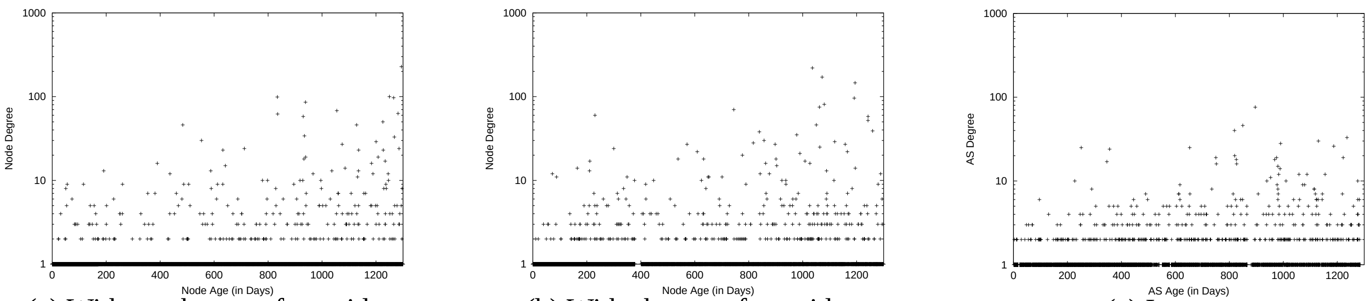
Unfortunately, inference for historical AS evolution (i.e., AS birth, AS death, change-of-provider, becoming a peer or customer AS, etc.) is compromised by the absence of reliable AS meta-data. This difficulty not withstanding, we rely on a hand-crafted approach for obtaining qualitative rather than quantitative evidence in support of including AS evolution into our multivariate HOT model. In particular, relying on the daily data sets from the Oregon route server which span the period Nov. 1997 to May 2001 and applying a previously reported methodology for identifying actual AS births (see [12]), we carefully extracted a set of ASs that were born during this period.[19] For those newly born ASs that were still alive at the end of our data collection period (May 26, 2001), we calculated their ages with respect to May 26, 2001. For the AS_{PC} subgraph inferred from the Internet AS topology of May 26, 2001, we obtained the actual ages of 8,799 ASs. The age difference between provider-customer pairs when both ASs belong to these 8,799 ASs can be easily computed. For ASs born before the start of Oregon's data collection effort (Nov. 1997) and whose exact age information is therefore not available, we set their ages to the maximum age of the 8,799 ASs.

Given an AS X, if all the current parent ASs of X are *younger* than X, X is considered to have undergone a change-of-provider. By this criterion, almost 3% of the ASs present in the May 26, 2001 provider-customer subgraph have changed provider at least once during their lifetime. With a less stringent definition of what constitutes a change-of-provider (i.e., a fraction of AS X's parent ASs are younger than X), some 15% of all the ASs present in the May 26, 2001 subgraph may have undergone a change of providers. As to age-degree correlation, plot (c) in Fig. 13 shows the scatter-plot for those 8,799 AS for which we have actual age information and which are part of the

[17] To generate a graph with N final nodes using the multivariate model, we need to add $N \cdot \frac{P_d}{P_d - 1}$ nodes, out of which $\frac{N}{P_d - 1}$ nodes will be removed during the growth process.

[18] The chosen value of P_d is an approximate ratio of AS birth frequency and AS death frequency that are empirically observed from the Oregon data sets described in Section VI-B.

[19] We also experimented with using AS number as a substitute for AS age, but found that due to AS number recycling and the specifics of AS number allocation (e.g., disjoint AS number ranges are allocated to the different Regional Internet Registries, which in turn assigned individual AS numbers sequentially), inferring AS age form AS number may be too crude of an approximation.



(a) Without change-of-provider (b) With change-of-provider (c) Internet

Fig. 13. Age-degree correlation

(single-homed) provider-customer subgraphs of the OREGON+ AS graph. As can be seen, there is no pronounced correlations between the two quantities; many of the older ASs have degree one or two, and the age of those ASs with degree bigger than, say, 10 ranges widely across the entire x-axis. While this plot agrees qualitatively with that of our multivariate HOT model in plot (b)—in that both suggest little correlation between AS age and AS node degree, they also show a discrepancy in the range of node degrees. We conjecture that the absence of 100+ degree nodes in plot (c) is due to the fact that in the real Internet, high-degree customer ASs will attempt to become peers of their providers. This aspect of AS evolution has been explicitly ignored by our deliberate focus on the provider-customer portion of the AS graph, but arises at this point as a natural next step in the development of a genuine HOT model for Internet growth at the AS level as a whole. The development of such a more complete HOT model that also accounts for peer ASs (i.e., their historical evolution and appropriate business models) and peer-to-peer relationships is beyond the scope of this paper and is left as future work.

VII. DISCUSSION AND OUTLOOK

One of the features of the proposed multivariate HOT model for Internet growth at the AS_{PC} level is that by explaining the striking power-law type node degree distributions of inferred AS graphs and subgraphs in terms of power-law type distributions for inferred AS size (i.e., the number of PoPs per AS), it confirms a previously reported conjecture in [31], namely that "*the highly variable degree distribution may arise merely from its correlation with a highly variable size distribution.*" Naturally, such an explanation can be rightly criticized as answering one question (i.e., what causes power-law type AS degree distributions?) with another question (i.e., what causes power-law type AS size distribution?) To respond to this criticism, first note that our explanation has shifted the focus from an abstract concept (i.e., node degree) to a concrete object (i.e., AS). Moreover, by viewing ASs as genuine businesses or firms, we can now bring to bear an extensive existing literature on empirical studies of firm size distributions that provide compelling evidence of the ubiquity of power-law type distributions of firm sizes, irrespective of how firm size is measured (e.g., number of employees, revenues, sales volume, customer base, number branch offices or outlets). In this sense, the situation is comparable to explaining the high variability or power-law type distributions exhibited by the constituent components of aggregate network traffic (e.g., sizes or durations of TCP connections, IP flows, macro-flows, etc.) in terms of power-law type distributions for the sizes of files on file servers or the sizes of Web documents on Web servers. The latter have proven to be a broadly robust properties over time and across different servers, despite often drastic changes in user behavior or application design. Similarly, firm sizes, measured variously, have been shown to consistently exhibit power-law type behavior, over many years and decades and for different nations, despite constant merger and acquisition activities, as well as bankruptcies (see for example [5], [27] and references therein). Thus our approach is to take the high variability of firm sizes for granted, just as we have come to accept the power-law type distributions for file or Web document sizes. However, we also want to point out that studies that attempt to explain the power-law type behavior of file sizes or firm sizes are of interest in their own right and have attracted considerable attention (e.g., [33], [14], [25], [4]).

A major contribution of this paper is make the AS business model concept a vital part of the multivariate HOT model for Internet growth at the AS_{PC} level. By abstracting the vague concept to an intuitively appealing but admittedly naive mathematical toy model, we demonstrated its potential for exploring some simple what-if scenarios through a systematic investigation of the model's low-dimensional parameterization. For example, our findings in Section V-C suggest that in today's Internet, the eco-

nomics of establishing provider-customer relationships is not very complex and appears to be based on a relatively small number of basic criteria and objectives. In turn, our toy AS business models also predict that if an ISP's primary task in the future is no longer simply to amass adequate resources to build out its infrastructure to fuel the Internet's overall growth, but becomes more sophisticated due to emerging refinements of interconnection business structures, we can expect to see qualitative changes in the resulting AS graph structures. However, a careful investigation of the full potential of the business model concept, including more sophisticated, elegant, and versatile abstractions is left for future research.

We emphasize that while the proposed multivariate HOT model has been designed to explicitly describe Internet growth at the AS_{PC} level and **not** at the overall AS level (i.e., including peer-to-peer relationships), it is nevertheless capable of explaining observed phenomena in inferred (overall) AS graphs, provided the essential characteristics of these phenomena are already present in the AS_{PC} subgraphs. The high variability of AS node degrees is one such characteristic. To develop a HOT model for Internet growth that applies to the entire AS graph and not just to the portion corresponding of the AS_{PC} subgraph would require an appropriate treatment of peer-to-peer relationships, either within the framework of our current AS business model or by modifying the present model to explicitly account for the possibilities of customer ASs becoming peers and peers turning into customers (see for example our conjecture in Section VI). In either case, the resulting HOT model can be expected to lend itself to easy generation of realistic topologies at the AS level, with the appealing property that Internet-like AS connectivity is obtained and guaranteed by imitating the very distributed and decentralized approach that underlies the design of AS connectivity in the actual Internet. All the model parameters have physical meaning, and using them as "knobs" results in predictable and intuitively easy-to-grasp refinements of the resulting graphs' overall structures.

Finally, within the broader context of Internet modeling as a science and in light of an extensive literature that advocates the orthodox physics view that power-law type behavior is unambiguously related to critical phase transition [6], our approach suggests a simple recipe for separating sound from specious claims and theories: use domain knowledge and check against appropriate measurements. For example, when comparing the scale-free models for Internet growth at the AS level introduced in [8], [2] with our multivariate HOT model, we find that the former is void of any domain knowledge and can be easily refuted using available measurements about the network's historical evolution [12], [32]. In contrast, the latter not only thrives on domain knowledge and incorporates it explicitly into the model formulation, but is also fully consistent with a number of available measurements that provide relevant information about all different aspects of the model. Given such an attractive alternative, it will be difficult to argue for the networking relevance of models and theories that can be easily refuted, both analytically as well as empirically.

REFERENCES

[1] W. Aiello, F. Chung, and L. Lu. A Random Graph Model for Massive Graphs. In *Proc. of ACM STOC*, 2000.

[2] R. Albert and A.-L. Barabási. Topology of Evolving Networks: Local Events and Universality. *Physical Review Letters*, 85(24):5234–5237, 2000.

[3] D. Alderson, J. Doyle, R. Govindan, and W. Willinger. Toward an Optimization-Driven Framework for Designing and Generating Realistic Internet Topologies. In *Proc. of HotNets-I*, 2002.

[4] R. Axtell. The emergence of firms in a population of agents: Local increasing returns, unstable Nash equilibria, and power law size distributions. Technical Report CSED Working Paper No.3, Brookings Inst., 2001.

[5] R. Axtell. Zipf distribution of U.S. firm sizes. *Science*, 293:1818–1820, 2001.

[6] P. Bak. *How Nature Works: The Science of Self-Organized Criticality*. Springer-Verlag, 1999.

[7] A.-L. Barabási. *Linked: The New Science of Networks*. Perseus Publishing, 2002.

[8] A.-L. Barabási and R. Albert. Emergence of Scaling in Random Networks. *Science*, 286:509–512, October 1999.

[9] A. Bookstein. Informetric distributions, Part I/II. *Journal of the Amer. Soc. for Information Science*, 41:368–386, 1990.

[10] CAIDA. NetGeo - The Internet Geographic Database. http://www.caida.org/tools/utilities/netgeo/.

[11] J. M. Carlson and J. Doyle. Highly Optimized Tolerance: A Mechanism for Power-Laws in Designed Systems. *Physical Review E*, 60(2):1412–1427, 1999.

[12] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. The Origin of Power Laws in Internet Topologies Revisited. In *Proc. of IEEE INFOCOM*, New York, NY, June 2002.

[13] J. Doyle and J. M. Carlson. Power Laws, Highly Optimized Tolerance and Generalized Source Coding. *Physical Review Letters*, 84(24):5656–5659, 2000.

[14] A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou. Heuristically Optimized Trade-offs. In *Proc. of ACM STOC*, 2002.

[15] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On Power-Law Relationships of the Internet Topology. In *Proc. of ACM SIGCOMM*, September 1999.

[16] A. M. Feldman. *Welfare Economics and Social Choice Theory*. Kluwer, Boston, 1980.

[17] L. Gao. On Inferring Autonomous System Relationships in the Internet. In *Proc. of IEEE Globecom*, San Francisco, CA, 2000.

[18] Genuity. U.S. Dedicated Access PoPs list. http://www.genuity.com.

[19] R. Govindan and P. Radoslavov. An Analysis of The Internal Structure of Large Autonomous Systems, 2002. unpublished manuscript.

[20] G. Huston. Analyzing the Internet BGP Routing Table. *The Internet Protocol Journal*, 4(1), Mar. 2001.

[21] C. Jin, Q. Chen, and S. Jamin. Inet: Internet Topology Generator. Technical Report CSE-TR-433-00, EECS Dept., University of Michigan, 2000.

[22] A. Lakhina, J. Byers, M. Crovella, and I. Matta. On the Geographic Location of Internet Resources. In *Proc. of ACM SIGCOMM Internet Measurement Workshop*, 2002.

[23] Level 3 Communication, Inc. Internet access customer buying guide, 2002. http://www.level3.com.

[24] A. Medina, A. Lakhina, I. Matta, and J. Byers. BRITE: An Approach to Universal Topology Generation. In *Proc. of MASCOTS '01*, August 2001.

[25] M. Mitzenmacher. Dynamic Models for File Sizes and Double Pareto Distributions, 2002. preprint.

[26] V. N. Padmanabhan and L. Subramanian. An Investigation of Geographic Mapping Techniques for Internet Hosts. In *Proc. of ACM SIGCOMM*, August 2001.

[27] J. Ramsden and G. Kiss-Haypal. Company size distribution in different countries. *Physica A*, 277:220–227, 2000.

[28] Route-Views. University of Oregon Route Views Project. http://www.routeviews.org.

[29] N. Spring, R. Mahajan, and D. Wetherall. Measuring ISP Topologies with Rocketfuel. In *Proc. of ACM SIGCOMM*, Pittsburgh, PA, 2002.

[30] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz. Characterizing the Internet Hierarchy from Multiple Vantage Points. In *Proc. of IEEE INFOCOM*, 2002.

[31] H. Tangmunarunkit, J. Doyle, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Does AS Size Determine Degree in AS Topology? In *ACM Computer Communication Review*, October 2001.

[32] W. Willinger, R. Govindan, S. Jamin, V. Paxson, and S. Shenker. Scaling phenomena in the Internet: Critically examining criticality. In *Proc. of the National Academy of Sciences*, 2001.

[33] X. Zhu, J. Yu, and J. Doyle. Heavy Tails, Generalized Coding, and Optimal Web Layout. In *Proc. of IEEE INFOCOM*, 2001.

APPENDIX

I. ON INFERRING AS BIRTH RATE

Fig. 14(a) and (b) plot daily AS birth rate (i.e., the number of ASs born per day) and the total number of ASs born since Nov. 1997, respectively. According to Fig. 14(a), AS birth rate has remained flat till day 600 or so, after which it started to ramp up till around day 800, and has been stable ever since. Fig. 14(b) confirms this AS birth rate transition. That is, the "total number of births" curve in Fig. 14(b) can be approximated by concatenating two linear lines that meet around day 700. Based on this observation, we characterize historical AS birth rate by two distinctly constant values, which is

$$birth\_rate(t) = \begin{cases} C_1, & \text{if } 0 < t < T_0; \\ C_2, & \text{if } t \geq T_0, \end{cases} \tag{1}$$

where $t = 0$ is the time when the Internet AS graph came into being, $t = T_0$ is the time when the transition of AS birth rate occurred, and C_1 and C_2 are positive constants with $C_1 < C_2$.[20]

Calculating the slopes of the curve of Fig. 14(b) in two separate time periods [0,700] and [700,1200] yields C_1 = 5.4274 and C_2 = 13.1. In order to approximately determine T_0, we note that:

# of ASs at T_0 = # of AS births during [0,T_0] - # of AS deaths during [0,T_0] = (1-$\frac{1}{P_d}$)$C_1 T_0$,

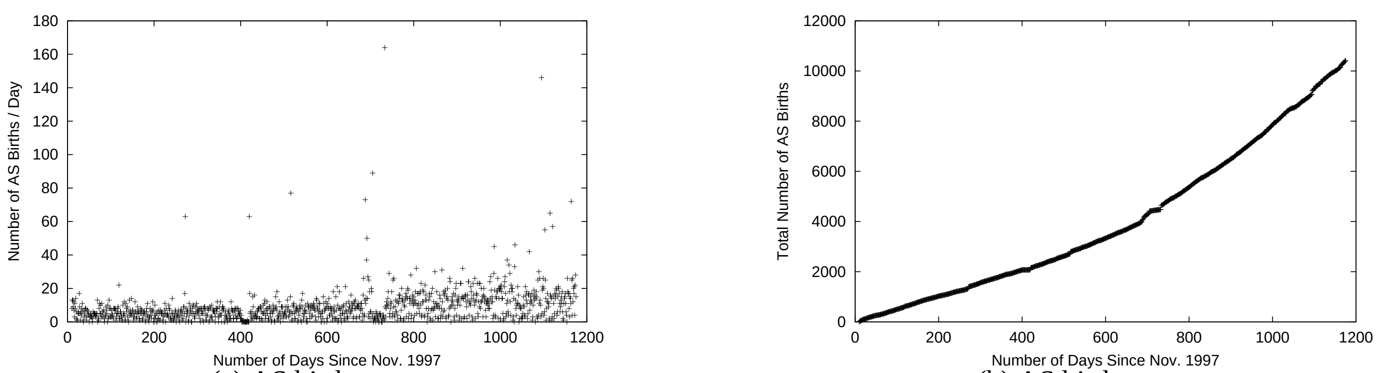
where P_d is the ratio of the number of AS births and the number of AS deaths during [0,T_0]. However, since the AS birth/death history data is available only from Nov. 1997, we approximate P_d with $\tilde{P}_d$, which is the ratio of the number of AS births and the number of AS deaths during [Nov. 1997,T_0]. This approximation yields T_0 = 1,390, which translates the birth time of the Internet to Jan. 1996.

If the node birth rate as shown in (1) is assumed in our multivariate model, the age of node i in a graph where the maximum node id is N can be characterized as:

$$age(i) = \begin{cases} (\frac{N}{C_2} - \frac{i}{C_1}) + \frac{C_2 - C_1}{C_2} \cdot T_0, & \text{if } i < C_1 T_0; \\ \frac{N-i}{C_2}, & \text{if } i > C_1 T_0. \end{cases} \tag{2}$$

Node ages calculated by the above formula have the same time unit as in the Internet AS graph (i.e., the number of days), and therefore node ages and AS ages can be directly compared.

[20]Note that the available historical AS data sets do not date back to t = 0. The start date (Nov. 1997) of the date sets lies between t=0 and t=T_0. For the purpose of this paper, we assume that AS birth rate for the period [0, Nov. 1997] has remained the same as that for the period [Nov. 1997, T_0].



(a) AS birth rate

(b) AS birth sum

Fig. 14. AS birth history